

Gesellschaft für Informatik (GI)

publishes this series in order to make available to a broad public recent findings in informatics (i.e. computer science and information systems), to document conferences that are organized in cooperation with GI and to publish the annual GI Award dissertation.

Broken down into the fields of

- Seminar
- Proceedings
- Dissertations
- Thematics

current topics are dealt with from the fields of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure the high level of the contributions.

The volumes are published in German or English

Information: <http://www.gi-ev.de/service/publikationen/lni/>

ISSN 1617-5468

ISBN 978-3-88579-232-1

This volume contains papers from the DW2008 Conference on Information Logistics, Integration and Service Orientation held in St. Gallen October 27-28, 2008. The topics covered in the papers range from advances in data warehousing and business intelligence, integration management and enterprise architecture to SOA and component orientation.



Barbara Dinter, Robert Winter, Peter Chamoni,
Norbert Gronau, Klaus Turowski (Hrsg.): DW2008

138

GI-Edition

Lecture Notes in Informatics

**Barbara Dinter, Robert Winter,
Peter Chamoni, Norbert Gronau,
Klaus Turowski (Hrsg.)**

Synergien durch Integration und Informationslogistik

**DW2008, 27.-28. Okt. 2008
St. Gallen, Schweiz**

Proceedings





Barbara Dinter, Robert Winter, Peter Chamoni,
Norbert Gronau, Klaus Turowski (Hrsg.)

DW2008

Synergien durch Integration und Informationslogistik

27./28.10.2008
in St. Gallen, Switzerland

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) – Proceedings

Series of the Gesellschaft für Informatik (GI)

Volume P-138

ISBN 978-3 88579-232-1

ISSN 1617-5468

Volume Editors

Barbara Dinter, Robert Winter

Universität St. Gallen, Institut für Wirtschaftsinformatik IWI-HSG, St. Gallen

Email: {barbara.dinter | robert.winter} @unisg.ch

Peter Chamoni

Universität Duisburg-Essen, MSM – Mercator School of Management, Duisburg

Email: peter.chamoni@uni-due.de

Norbert Gronau

Universität Potsdam, Lehrstuhl für Wirtschaftsinformatik und Electronic

Government, Potsdam

Email: norbert.gronau@wi.uni-potsdam.de

Klaus Turowski

Universität Augsburg, Institut für Wirtschaftsinformatik und Systems Engineering,
Augsburg

Email: klaus.turowski@wiwi.uni-augsburg.de

Series Editorial Board

Heinrich C. Mayr, Universität Klagenfurt, Austria (Chairman, mayr@ifit.uni-klu.ac.at)

Jörg Becker, Universität Münster, Germany

Hinrich Bonin, Leuphana-Universität Lüneburg, Germany

Dieter Fellner, Technische Universität Darmstadt, Germany

Ulrich Flegel, SAP Research, Germany

Johann-Christoph Freytag, Humboldt-Universität Berlin, Germany

Ulrich Furbach, Universität Koblenz, Germany

Michael Koch, TU München, Germany

Axel Lehmann, Universität der Bundeswehr München, Germany

Peter Liggesmeyer, TU Kaiserslautern und Fraunhofer IESE, Germany

Ernst W. Mayr, Technische Universität München, Germany

Heinrich Müller, Universität Dortmund, Germany

Sigrid Schubert, Universität Siegen, Germany

Martin Warnke, Leuphana-Universität Lüneburg, Germany

Dissertations

Dorothea Wagner, Universität Karlsruhe, Germany

Seminars

Reinhard Wilhelm, Universität des Saarlandes, Germany

Thematics

Andreas Oberweis, Universität Karlsruhe (TH)

© Gesellschaft für Informatik, Bonn 2008

printed by Köllen Druck+Verlag GmbH, Bonn

Vorwort und Einführung

Mehr denn je stehen Unternehmen vor der Herausforderung, flexibel und zeitnah auf sich ändernde Rahmenbedingungen und daraus resultierende neue fachliche oder technische Anforderungen zu reagieren. Dieser Anspruch beeinflusst auch in hohem Maße die betriebliche Informationsverarbeitung und -versorgung. Nicht nur die unterstützenden Informationssysteme gilt es entsprechend anzupassen bzw. zu erweitern, auch gewinnt die Integration zunehmend an Bedeutung. Integration ist hierbei mehrschichtig zu sehen: Zum einen wird die technische Integration auf Daten- und Applikationsebene betrachtet. Zum anderen gilt es, die Integration und den Abgleich zwischen fachlichen und technischen Strukturen zu unterstützen. Zudem rücken analytische Informationssysteme und operative Applikationen immer stärker zusammen. Erst ihre Verknüpfung stellt die zeitnahe und vollständige Bereitstellung von Informationen zur Entscheidungsunterstützung im Unternehmen sicher und kann die Wettbewerbsfähigkeit effektiv unterstützen.

Die Konferenz DW2008 adressiert diese Herausforderungen in drei Tracks. Sie spannt den Bogen von Datenintegration (Informationslogistik und Data Warehousing) hin zur Applikations-, Service- und Prozessintegration (Enterprise Application Integration, serviceorientierte Architekturen). Einen weiteren Schwerpunkt bildet die bedarfsgerechte Gestaltung von Unternehmens- und Softwarearchitekturen. Die DW2008 ist die bereits fünfte Veranstaltung der DW-Reihe, die seit 2000 im Zweijahres-Rhythmus stattfindet.

Der *Track Informationslogistik/Data Warehousing und Datenqualitätsmanagement* adressiert die Bereitstellung hochqualitativer analytischer Informationen für Entscheider und Prozesse. Im Mittelpunkt stehen die Konzeption, Realisierung und Nutzung analyseorientierter Systeme, aber auch neue Entwicklungen wie beispielsweise die integrierte Betrachtung von Business Intelligence und Geschäftsprozessmanagement. In den Beiträgen des Tracks wird der Bogen gespannt von Strategie und Governance der Informationslogistik über Fragestellungen der Architekturgestaltung und des Betriebs von Informationslogistik-Systemen bzw. Infrastrukturthemen wie Datenqualitäts-, Metadaten- und Stammdatenmanagement bis hin zu Praxisberichten.

Der *Track Integration und Integrationsmanagement* konzentriert sich auf die bedarfsrechte Konzeption, Realisierung und Nutzung integrierter Anwendungssysteme. Die Beiträge im Track adressieren Themen des Architekturmanagements ebenso wie der Integration von Anwendungen, der Governance serviceorientierter Architekturen (SOA) und der Datenmigration.

Der *Track Komponenten- und Serviceorientierung* behandelt die modulare Gestaltung von Unternehmensarchitekturen und speziell Anwendungssoftwarearchitekturen. Im Mittelpunkt steht die Frage, wie Komponenten- und Serviceorientierung helfen können, über Konzepte wie Wiederverwendung und Interoperabilität Unternehmensarchitekturen zu flexibilisieren und ihre Agilität zu steigern. Die Beiträge des Tracks haben ausgewählte Fragestellungen der Komponentenorientierung und serviceorientierter Architekturen zum Inhalt und stellen SOA-Praxisbeispiele vor.

Die in diesem Band zusammengefassten wissenschaftlichen Beiträge und Praxisberichte werden auf der Konferenz „DW2008 – Synergien durch Integration und Informationslogistik“ am 27. und 28. Oktober 2008 an der Universität St. Gallen präsentiert. Zur Konferenz wurden 68 Beiträge eingereicht, die das Programmkomitee in einem Doppelt-Blind-Verfahren begutachtete. Insgesamt wurden 32 Beiträge angenommen (Annahmequote 47 %). Da im Aufruf zu Einreichungen Wissenschaft und Praxis gleichermaßen angesprochen wurden, finden sich im Konferenzband sowohl Ergebnisse aus Forschungsaktivitäten wie auch Erkenntnisse aus Praxisprojekten.

Konferenzen und Herausgeberbände sind immer das Ergebnis eines ideell motivierten Engagements, für das die Herausgeber insbesondere den Mitgliedern des Programmkomitees, die alle eingegangenen Beiträge begutachteten, sowie natürlich den Autoren herzlich danken. Ein besonderer Dank gilt auch den Hauptsponsoren IBM und hp/Microsoft Frontline Partnership sowie allen weiteren Sponsoren, die mit ihrer finanziellen Beteiligung die Durchführbarkeit dieser Konferenz sicher stellen.

Unser Dank gebührt insbesondere den Personen, die an der Organisation der Konferenz sowie an der Erstellung des Konferenzbands unmittelbar mitgewirkt haben. Zu erwähnen sind hier insbesondere Dr. Stephan Aier, Bettina Gleichauf und Philipp Gubler sowie alle anderen beteiligten wissenschaftlichen und studentischen Mitarbeitern/innen am Institut für Wirtschaftsinformatik der Universität St. Gallen.

Die Herausgeber

St. Gallen, im August 2008

Für die breite Einbindung der wissenschaftlichen Gemeinde sowie die Unterstützung bei der inhaltlichen Gestaltung der Tracks und der Begutachtung der Einreichungen ist der tatkräftigen Unterstützung der Track-Chairs Prof. Dr. Peter Chamoni (Track Informationslogistik), Prof. Dr. Norbert Gronau (Track Integration) und Prof. Dr. Klaus Turowski (Track Komponenten- und Serviceorientierung) und ihren beteiligten Mitarbeitern/innen großer Dank geschuldet. Schließlich danken wir den Herausgebern der LNI-Reihe für die Aufnahme des Konferenzbandes in diese Reihe, Cornelia Winter aus der Geschäftsstelle der Gesellschaft für Informatik sowie dem Team vom Köllen-Verlag für die angenehme und professionelle Zusammenarbeit.

Barbara Dinter, Robert Winter

St. Gallen, im August 2008

DW2008 – Synergien durch Integration und Informationslogistik

Veranstalter

Universität St. Gallen
Institut für Wirtschaftsinformatik

Prof. Dr. Robert Winter (Vorsitzender des Programmkomitees)
Dr. Barbara Dinter (Wissenschaftliches Programm)
Dr. Stephan Aier (Ausstellung)
Bettina Gleichauf, Philipp Gubler (Organisation)

Programmkomitee

S. Aier, Universität St. Gallen
A. Albani, Universität Augsburg/TU Delft
C. Atkinson, Universität Mannheim
C. Bange, BARC
W. Behme, Continental AG
A. Bernstein, Universität Zürich
J. vom Brocke, Hochschule Liechtenstein
U. Bub, Deutsche Telekom Laboratories
P. Chamoni, Universität Duisburg-Essen (Vorsitz Track Informationslogistik)
P. Dadam, Universität Ulm
B. Dinter, Universität St. Gallen
C. Felden, TU Bergakademie Freiberg
R. Flatscher, Wirtschaftsuniversität Wien
U. Frank, Universität Duisburg-Essen
R. Gabriel, Universität Bochum
M. Gaedke, Technische Universität Chemnitz
P. Gluchowski, Technische Universität Chemnitz
N. Gronau, Universität Potsdam (Vorsitz Track Integration)
O. Höft, O2 Germany
T. Hübsch, Deutsche Telekom
R. Jung, Universität Duisburg-Essen
F. Kaminsky, Zürich Versicherung
H.-G. Kemper, Universität Stuttgart
K. Labis, Zürcher Kantonalbank
P. Loos, Universität Saarbrücken
J. Lüssem, HSH Nordbank

W. Martin, Wolfgang Martin Team
M. Meier, Universität Augsburg
D. Meyer, E.ON AG
F. Naumann, Hasso-Plattner-Institut
V. Nissen, Technische Universität Ilmenau
E. Ortner, Technische Universität Darmstadt
B. Otto, Universität St. Gallen
S. Overhage, Universität Augsburg
W. Pree, Universität Salzburg
W. Pussak, Bayer Business Services
C. Rautenstrauch, Universität Magdeburg
J. Rick, Deutsche Bahn AG
B. Rieger, Universität Osnabrück
J. Schelp, Universität St. Gallen
M. Schönherr, Deutsche Telekom Laboratories
E. J. Sinz, Universität Bamberg
C. Szyperski, MS Research/Queensland University
K. Turowski, Universität Augsburg (Vorsitz Track Komponenten- und Serviceorientierung)
R. Weinreich, Universität Linz
K. Weronek, Fraport AG
M. Weske, Hasso-Plattner-Institut
R. Winter, Universität St. Gallen (Vorsitz DW2008)

Hauptsponsorenverzeichnis



IBM Schweiz AG

Ansprechpartner: Herr Daniel Dalle Carbonare
Vulkanstrasse 106
CH-8010 Zürich
Tel.: +41 58 333 69 26
Fax: +41 58 333 40 40
Email: daniel.carbonare@ch.ibm.com
URL: <http://www.ibm.ch>



Microsoft

FRONTLINE PARTNERSHIP

HP Microsoft Frontline Partnership

Ansprechpartner: Herr Dr. Bernd Kiupel
Richtstrasse 3
CH-8304 Wallisellen
Tel.: +41 78 844 63 88
Fax: +41 43 456 44 44
Email: berndki@microsoft.com
URL: <http://www.microsoft.ch>

Sponsorenverzeichnis



Ab Initio Software Germany GmbH

Ansprechpartner: Herr Anthony Ibrahim
Landsberger Strasse 302
D-80687 München
Tel.: +49 89 90405 801
Fax: +49 89 90405 809
Email: aibrahim@abinitio.com
URL: <http://www.abinitio.com>



business solution group TI

Ansprechpartner: Herr Thomas Wüst
Buckhauserstrasse 24
CH-8048 Zürich
Tel.: +41 44 497 77 77
Fax: +41 44 497 77 70
Email: thomas.wuest@bsgroup.ch
URL: <http://www.bsgroup.ch>



IDS Scheer Schweiz AG

Ansprechpartnerin: Frau Concetta Saia
Industriestrasse 50a
CH-8304 Wallisellen
Tel.: +41 58 958 99 00
Fax: +41 58 958 97 01
Email: info-ch@ids-scheer.com
URL: <http://www.ids-scheer.ch>



saracus consulting AG

Ansprechpartnerin: Frau Daniela Moss
Täferstrasse 4
CH-5405 Baden-Dättwil
Tel.: +41 56 483 02 20
Fax: +41 56 483 02 21
Email: d.moss@saracus.com
URL: <http://www.saracus.com>



sd&m Schweiz AG

Ansprechpartner: Herr Dr. Holger Rommel
Leutschenbachstrasse 95
CH-8050 Zürich
Tel.: +41 44 307 83 00
Fax: +41 44 307 83 01
Email: holger.rommel@sdm.ch
URL: <http://www.sdm.ch>

SHS VIVEON AG

SHS VIVEON AG

Ansprechpartnerin: Frau Daniela Schneider
Fraunhoferstr. 12

D-82152 Martinsried b. München

Tel.: +49 8974 7257 286

Fax: +49 8974 7257 10

Email: daniela.schneider@shs-viveon.com

URL: <http://www.shs-viveon.com>



Steria Mummert Consulting AG

Ansprechpartner: Herr Klaus-Dieter Schulze
Hans-Henny-Jahnn Weg 29

D-22085 Hamburg

Tel.: +49 40 22703 0

Fax: +49 40 22703 7900

Email: BI@steria-mummert.de

URL: <http://www.steria-mummert.de>



Teradata Schweiz GmbH

Ansprechpartner: Business Support

Postfach

CH-8301 Glattzentrum

Tel.: +41 44 832 15 15

Fax: +41 44 832 15 51

Email: tb430007@teradata.com

URL: <http://www.teradata.com>



Trivadis Holding AG

Ansprechpartnerin: Frau Julia Baier

Europa-Strasse 5

CH-8152 Glattbrugg

Tel.: +41 44 808 70 20

Fax: +41 44 808 70 21

Email: info@trivadis.com

URL: <http://www.trivadis.com>

Medienpartner



BI-SPEKTRUM – Fachzeitschrift für Business Intelligence und Data Warehousing

Ansprechpartnerin: Frau Susanne Herl

Lindlastr. 2c

D-53842 Troisdorf

Tel.: +49 2241 2341 550

Fax: +49 2241 2341 199

E-Mail: susanne.herl@sigs-datacom.de

URL: <http://www.bi-spektrum.de>



is report – OXYGON Verlag GmbH

Ansprechpartner: Herr Stefan Raupach

Herzog-Heinrich-Str. 5

D-80336 München

Tel.: +49 89 904862 30

Fax: +49 89 904862 55

E-Mail: sraupach@oxygen.de

URL: <http://www.isreport.de>

Inhaltsverzeichnis

Vorwort und Einführung	5
------------------------------	---

Informationslogistik/Data Warehousing und Datenqualitätsmanagement

A framework for efficient information management <i>M. Dreher, D. Goetze, R. Wieland, D. Liebhart</i>	19
--	----

BI-Governance – Königsdisziplin strategischer Business Intelligence <i>U. Trost, M. Zirkel</i>	39
---	----

Das IT Performance Measurement Maturity Model – Ein Reifegradmodell für die Business Intelligence-Unterstützung des IT-Managements <i>J. Becker, R. Knackstedt, J. Pöppelbuß, L. Schwarze</i>	53
---	----

Bereitstellung und Nutzung von Endbenutzungs-Metadaten im Kontext der Informationslogistik: Empirische Erkenntnisse und Gestaltungsempfehlungen <i>T. Bucher, U. Wlk</i>	75
--	----

Proaktives Management von Konsistenzbedingungen im Analytischen Performance Management <i>S. Brüggemann</i>	95
---	----

Text-Mining-Einsatz zur Bewertung der Reputation von Organisationen und Unternehmen <i>P. Kirchberg</i>	113
---	-----

Ein Konzept zur automatisierten Klassifizierung von Informationen für das Information Lifecycle Management – dargestellt am Beispiel des SAP NetWeaver Business Intelligence <i>M. Kosler, M. Matthesius, D. Stelzer</i>	129
---	-----

Harmonizing company-wide Information Objects <i>A. Schmidt, B. Otto</i>	147
--	-----

Ein Plädoyer für die Berücksichtigung von Semantik beim Stammdaten- Alignment – Vorgehensmodell und prototypische Anwendung im Einzelhandel <i>A. Winkelmann, M. Matzner, O. Müller, J. Becker</i>	165
--	-----

Effektive Architekturgestaltung auf Basis einer Typologie für Datenintegrationsarchitekturen <i>R. Jung</i>	183
---	-----

Management von Business Intelligence Services <i>T. Horakh, H. Baars, H.-G. Kemper</i>	199
---	-----

Business-Intelligence-Konzept auf Basis einer Event-Driven Service-Oriented Architecture <i>T. Vogt, S. Neuhaus, M. Linden, P. Chamoni</i>	217
Business Intelligence Services bei T-Mobile Deutschland: Service Level Agreements und servicebezogenes Datenqualitätsmanagement zur kundengerechten Leistungserbringung <i>C. Hermann, S.-A. Müller</i>	229
Governance in der Informationslogistik am Beispiel eines Energieversorgers <i>B. Dinter, G. Lahrman, D. Meyer, M. Schmaltz</i>	249
BI-Bebauungsplanung im Rahmen der BI-Strategie des Volkswagen-Konzerns <i>T. Sommer, W. Bayer, S. Rosenbaum, B. Wendland, A. Overmeyer</i>	267
Intelligente Wohnungsmarktanalyse auf Basis einer Business Intelligence-Anwendung <i>R. Gabriel, T. Hoppe, A. Pastwa</i>	287
Ein mandantenfähiges Data-Warehouse-System für die bayerischen Fachhochschulen – Anforderungen, Konzeption und Realisierung <i>T.-O. Förtsch, S. Hartmann, C. Jürck, J. Ringler</i>	299
Entwicklung eines Management Cockpits mit SAP BI 7.0: ein Erfahrungsbericht <i>J. Raimann, S. Schnee, N. Gudelj, T. Broda</i>	315
Das Ende der Magersucht in der Managementinformation <i>N. Bissantz, R. Zimmermann</i>	329
Organising Accountabilities for Data Quality Management – A Data Governance Case Study <i>K. Weber, L. Cheong, B. Otto, V. Chang</i>	347
Integration und Integrationsmanagement	
Szenariobasierte Entwicklung der Unternehmensarchitektur bei der Real-Time Center AG <i>S. Aier, N. Buchs, J. Saat, G. Steiner</i>	363
Integrationsmanagement mittels Quality Gates <i>T. Snaikus, G. Disterer</i>	379
Data Migration Project Management and Standard Software – Experiences in Avaloq Implementation Projects <i>K. Haller</i>	391
Ein Framework zur Einführung einer SOA-Governance in Unternehmen <i>C. Schröpfer, M. Schönherr</i>	407

Integration öffentlicher Anwendungslandschaften <i>N. Gronau, M. Stein, S. Eggert, A. Lämmer</i>	425
Aufbau und Einsatz der Geschäftsarchitektur bei der AXA Winterthur <i>S. Aier, F. Maletta, C. Riege, K. Stucki, A. Frank</i>	437
Ein Annotationsansatz zur Unterstützung einer ganzheitlichen Geschäftsanalyse <i>S. Radeschütz, F. Niedermann, B. Mitschang</i>	455
Komponenten- und Serviceorientierung	
Eine Serviceorientierte Architektur (SOA) für die Integration von RFID-, Sensor- und Lokalisierungsdaten im Facility Management <i>D. Hanhart, C. Legner, H. Österle</i>	475
Evaluation von Konzepten zur Integration von UDDI-Registries unter Berücksichtigung von Marktaspekten <i>N. Brehm, L. Haak, J. Marx Gómez</i>	491
Zentrale Gegenstände der Parametrisierung bei betrieblichen Softwarekomponenten <i>J. Ackermann, K. Turowski</i>	509
Serviceorientierte Architekturen – Potentiale für eine nachhaltige Steigerung der Unternehmensagilität <i>J. Schelp, S. Aier</i>	529
SOA Implementierung in der Luftfahrtindustrie am Beispiel der Lufthansa <i>R. Kubli, M. Burkard, E. Rantakallio, A. Kreczy</i>	543

Informationslogistik/Data Warehousing und Datenqualitätsmanagement

Leitung und Organisation des Tracks

P. Chamoni, Universität Duisburg-Essen

Programmkomitee des Tracks

C. Bange, BARC
W. Behme, Continental AG
P. Chamoni, Universität Duisburg-Essen
B. Dinter, Universität St. Gallen
C. Felden, TU Bergakademie Freiberg
R. Gabriel, Universität Bochum
P. Gluchowski, TU Chemnitz
R. Jung, Universität Duisburg-Essen
F. Kaminsky, Zürich Versicherung
H.-G. Kemper, Universität Stuttgart
K. Labis, Zürcher Kantonalbank
J. Lüssem, HSH Nordbank
W. Martin, Wolfgang Martin Team
M. Meier, Universität Augsburg
D. Meyer, E.ON AG
F. Naumann, Hasso-Plattner-Institut
B. Otto, Universität St. Gallen
B. Rieger, Universität Osnabrück

A framework for efficient information management

Martin Dreher*, Daniel Goetze**, Robert Wieland***, Daniel Liebhart****

* Prozess- und Informationsmanagement
E.ON Energy Sales GmbH
Karlstraße 68
80335 München
martin.dreher@eon-energie.com

** iteratec GmbH
daniel.goetze@iteratec.de

*** IBSolution GmbH
robert.wieland@ibsolution.de

**** Trivadis GmbH
daniel.liebhart@trivadis.com

Abstract: Since information is more and more recognised as a business relevant asset the need for an integrated information management approach grows steadily. Against this background large industrial companies begin to think about valid information governances and policies. In the following article the information management approach developed at E.ON is presented. It bases on a role concept, a meta-information repository and generic models for information quality management and information lifecycle management. It finally allows for an efficient implementation of the currently most important information management issues which are information quality and information lifecycle management.

1 Introduction

The spectacular Enron scandal in 2002 as well as the world-wide recognizable impact of the us-american sub-prime crisis with their far reaching effects are only two examples which show that issues like legal compliance, transparency or information quality are of great importance for all kinds of companies. This situation grows more severe since the need for high quality information grows and information continuously develops towards a business relevant asset. Against this background, companies must successfully cope with numerous legal requirements, like the Sarbanes-Oxley-Act or the European MiFID¹. In order to do so, a number of different governances and policies is developed

¹ Markets in Financial Instruments Directive

concerning mainly information-related topics like compliance to legal frameworks, information quality, and security. Typically, those governance or policy documents focus on requirements towards an information management. Issues about the implementation, like organizational frameworks or methods, are hardly addressed, although this is crucial for a successful realization of a governance or policy.

In this paper we want to introduce a concept for an effective and efficient information management, which was developed at E.ON. It is mainly based on an organizational framework described in chapter 2, a meta-information repository (see chapter 3) and a couple of generic models for the efficient handling of the currently most important issues of information quality and information lifecycle, which are presented in chapters 4 and 5. The document concludes with a brief look at the current state of implementation at E.ON (chapter 6) and a summary.

2 Organizational framework

For the implementation of an information governance the establishment of a corresponding organizational framework is a necessary base. Without such a frame information governance cannot be brought forward. This framework has an enabling function for some major drivers for information integration which are master data management, business intelligence, and service oriented architecture. Vice versa, these drivers support its implementation, too. In this context, it is necessary that the meaning of the information to be integrated is clearly defined and understood all through the enterprise. Furthermore, the information quality must be reliable and the information lifecycle must be defined. These common issues are addressed by the topics covered by information governance:

- *Information architecture management*: Leads to unambiguously defined information and a common understanding of how information currently is and in future will be used across the enterprise. It is tightly coupled with the other disciplines of enterprise architecture management: business and application architecture management.
- *Information quality management*: Gathers systematically the requirements concerning availability and reliability of information. Gives guidelines on how to measure and report information quality and how a fulfillment of requirements can be approved. Addresses information quality risks.
- *Information lifecycle management*: Gathers systematically the requirements concerning information lifecycle. Controls the fulfillment of the requirements to reduce the costs for the provisioning of information.

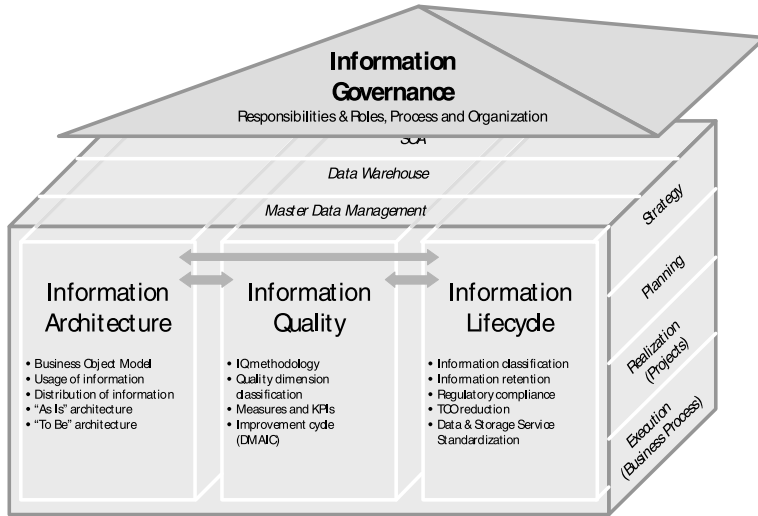


Figure 1: Organizational framework

2.1 Process Levels

Information governance is not established through stand alone, newly defined processes which solely address the needs of information. It rather should be ensured that (usually existing) processes are enriched in a way that awareness for information and adherence to guidelines and policies will be achieved. Information shall become a dedicatedly handled topic in these processes. Thus, the framework addresses "process levels" rather than "processes" and defines tasks to be fulfilled at these levels by the roles as described later on. Information governance becomes relevant at the levels of:

- *Strategy development:* At this level, the company's strategy derived from top level goals and the roadmap to achieve these goals are defined. This process level also subsumes the development of guidelines and policies, the setting of targets and the establishment of organizational changes (roles, processes).
- *Business- and IT-planning:* At this level e.g. demand management, mid-term planning and project portfolio management take place. Tasks are identified which are derived from strategic goals and business demands, prioritized and bundled to projects.
- *Realization of IT systems and organizational changes in projects:* A project is defined as any action that significantly changes business processes and/or IT systems. It is assumed that nearly every change is a result of a project. Size and duration of projects may vary significantly, but the basic organizational principles and rules stay the same. At E.ON, the project management guideline defines the underlying process.

- *Execution of business processes*: The day-to-day business performed at business- and IT-side.

2.2 Roles

The governance framework defines the roles that act in the context of the process levels described above. A role is a bundle of competencies, responsibilities and tasks that must be fulfilled to incorporate “information governance”. A role can be assigned to one single person or to a group of persons. On the other hand, one person may have several roles assigned. It is not implied that every role must be assigned to one employee. A coupling with other, possibly existing roles is possible and for the steward-roles even recommended.

Strong management support is needed to lead the organization in a direction that in fact handles information as an asset. Defining responsibilities and roles, the placement of roles in an enterprise’s organization and a clear commitment that governance related tasks must be fulfilled (and the fulfillment is proven) is crucial to bring information governance to life.

The *information governance board* with members out of the company’s top management provides the management attention. They are able to drive the decisions necessary to incorporate information governance in a reasonable way which takes into account the balance of efforts vs. results. One task is the final confirmation of information strategy, target information architecture, policies and guidelines as prepared by the information architect. It is up to the information architect to thoroughly prepare the decisions to be made by the board. In case of conflict, the board may be consulted as a last level of escalation.

The *information architect* is accountable for the current and future information architecture and the definition of strategic guidelines and policies. He is the head of a (virtual) team of information managers.

The *information manager* is a specialist in the field of one or more business objects and accountable for the fulfillment of overall requirements concerning the objects he is responsible for across the company’s processes and/or applications. He ensures reliable and consistent documentation of the information architecture and delivers operational support to resolve information quality issues together with the stewards.

The *business information steward*, as a (business) specialist in the field of one or more business processes, is accountable for gathering information related business requirements and adherence to information governance guidelines and policies in his area of competence. Often this role will additionally be assigned to a role like “process owner”.

The *technical information steward*, as a (technical) specialist in the field of one or more applications, is accountable for the fulfillment of requirements concerning “information” and adherence to information governance guidelines and policies in his area of competence. Often this role will additionally be assigned to an application specialist.

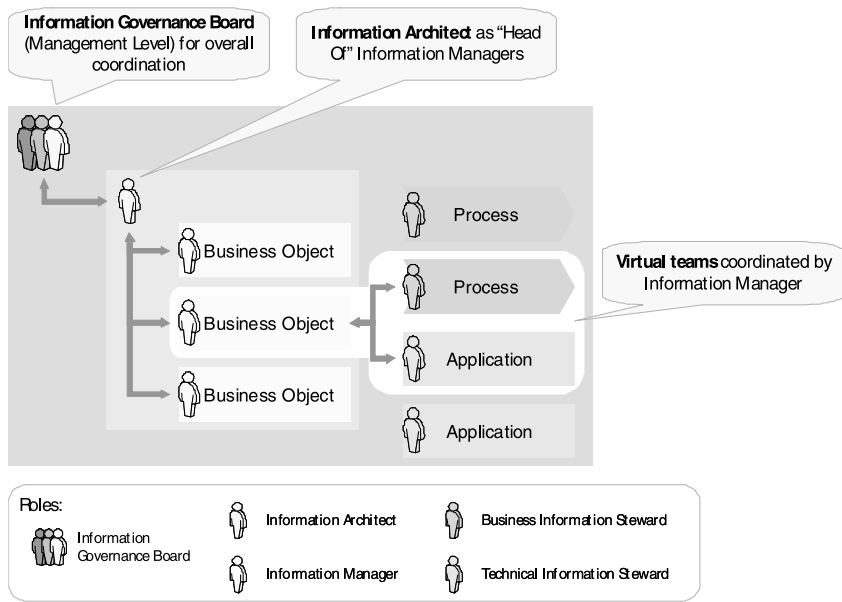


Figure 2: Cooperation between roles

The organizational link between information, processes and applications on an operational level is achieved by involving the steward roles in the governance organization. To be able to communicate, they need a common understanding of the things they are talking about. A meta-information repository helps to reach this common understanding and to systematically document the information items and their relation to processes and applications at different levels of abstraction. It serves as a public catalog of the objects affected by information governance, based on a committed standard for documentation.

3 The meta-information repository

The backbone for information governance activities is the repository which contains the documentation of the enterprise architecture. The enterprise architecture describes the objects affected by governance activities as well as their current and future state. As an outcome of governance activities, high-level classifications and requirements for the elements of the enterprise architecture are documented systematically and consistently. The elements serve as a container for a structured, traceable and reportable documentation that is centrally stored. Additionally, the person(s) responsible for a single item are documented.

The meta-model shown in *Figure 3* provides a coarse overview of the elements of the enterprise architecture and their relationships. Business processes, applications and application releases are considered in the context of the business- and application architecture.

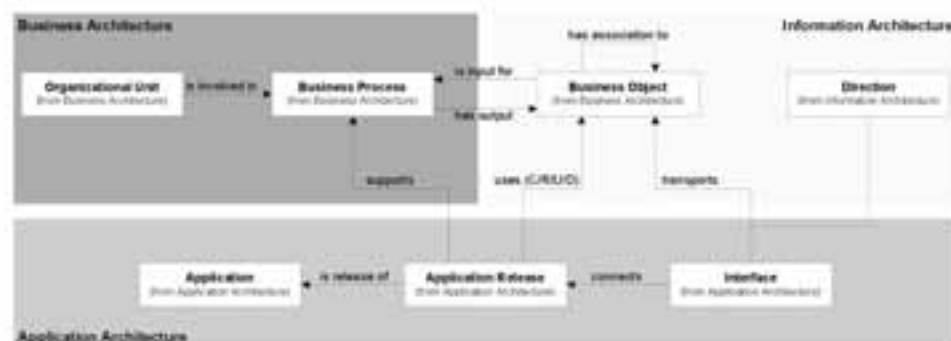


Figure 3: Enterprise Architecture meta-model (simplified)

The **business object** (BO) is the core element of an information architecture. Business objects are informational items which have major significance in the context of the business activity of a company. The primary question is: Which information items are relevant within the company, and which relationships exist between these items. On a BO level, no distinction is made between master data, transactional data, analytical data or metadata. Candidates for high level business objects are e.g. "product", "invoice" or "customer" and their respective specializations².

3.1 Business objects and their Context – the BO Profile

The BO-Profile shows a business object in its whole context, e.g. processes and applications using this business object. The profile can basically be used in two ways:

- It can serve as a template for manually specifying a business object, e.g. in the context of requirements management.
- It can be specified as a report in a specific modeling environment, like ARIS.

<i>BO Profile for <Name></i>	
<i>Description</i>	<Description/Definition>
<i>Contact persons</i>	Information Manager: <Name > Modeler: <Name >
<i>Associations to other business objects</i>	⇨ <association role> <Name of business object> ... (one line for every outgoing associations) ⇨ < Name of business object > <association role> ... (one line for every ingoing associations)

² If meta-information is modeled within a business object model, a "Business Object" is a business object itself.

BO Profile for <Name>	
<i>Business Value Class</i>	<Business Value Class>
<i>Lifecycle rules</i>	Operational (following state: analytical); <operational rule(s)>
	Analytical (following state: archived); <analytical rule(s)>
	Archived (following state: removed); <archived rule(s)>
<i>Using processes (incl. quality gates)</i>	<Name of business process>; <In/Out>, QGC: <quality gate class> ... (one line for every process using the object)
<i>Using application releases</i>	<Name of application> Type: <C/R/U/D> ... (one line for every application release using the object)
<i>Information flows</i>	<Name of application 1> ⇔ <Name of application 2> ... (one line for every information flow using the object)

Name and Description

Name and description of the business object. Under the given name you could find the object within an enterprise wide glossary.

Contact Persons

Information manager responsible for the business object and the modeler who modeled the object. Both may be the same person.

Associations to other business objects

A business object may have one or more directed associations with other business objects. For every association the role of the association and the associated business object are stated. E.g., in the BO-Profile for the business object itself one would see the associations³

- ⇔ Business Object *is input for* Business Process
- ⇔ Business Object *has association to* Business Object
- ⇔ Application Release *uses* Business Object
- ⇔ Business Process *has output* Business Object
- ⇔ Business Object *has association to* Business Object
- ⇔ Interface *transports* Business Object

³ see also Figure 1

Business Value Class

Assignment of a value-class to business objects from a business perspective. Business objects are rated from “*not important to operations*” to “*mission critical information*” according to their contribution to business operation (see chapter 5.3).

Lifecycle Rules

Rules that apply to the business object to trigger the end of one of the lifecycle phases⁴. The rule defines under which condition(s) the next phase starts. A rule is defined in the form

IF <condition> THEN following lifecycle state begins

<condition>: any condition (or combination of conditions) in human readable terms.

Using processes

Shows the processes that are using the business object and whether the business object is input for a process, output of a process, or both (e.g. if the object is changed during this process). Additionally, the “*quality gate class*” (QGC) of the using business process is shown (see chapter 4.2).

Putting together the *quality gate class* of the using process and the *business value class* of the business object itself, the quality dimensions and quality thresholds can be derived that are applicable within the respective business process (see chapter 4). This information allows to deduce which quality checks should be applied to the business object. Especially for business critical information, at least one process should be stated which validates the information quality of the business object.

Using application releases/information flows

Shows the using application releases and the information flow (incl. direction of the flow) between application releases. It is also defined how the business object is used within one application release. Possible values are (in any combination):

- C: Information is newly created
- R: Information is read
- U: Information is updated
- D: Information is deleted

⁴ The lifecycle phases and the lifecycle rules are defined in the context of information lifecycle management, see chapter 5.

3.2 The costs behind a BO-Profile

In a BO-Profile no “new” meta-information is documented. While designing a business process or writing a system proposal, the questions leading to the documented facts are normally addressed anyway. Therefore, there is little additional effort to gather and document this information in a structured way. The most effort lies in defining the structure (meta-model) of the repository, training people using this structure, gaining acceptance for documenting systematically, keeping the gathered information correct and consistent and, at last, deciding which objects to put into the repository to keep it manageable and on the right level of abstraction. Since at least for the beginning, the BO-Profiles can be maintained in a spreadsheet-table the initial system requirements are not very high. But it has to be considered that with an increasing number of profiles a simple spreadsheet-based solution quickly reaches its limits. Then, a professional tool for information architecture should be used.

4 Information Quality Management

This chapter introduces a generic model for **information quality measurement** that aligns the information quality analysis processes with business processes [Lo01]. To achieve the alignment between business processes and quality measures one has to know

- where to measure – the business processes
- what to measure – the relevant information quality attributes
- measure thresholds – which quality is “adequate” on dependence of the business value class of the business object.

The generic model enforces to think about information quality proactively at a time of process design. As a result, business processes obtain high(er) information quality, better project cycles (projects in budget) and process maintenance costs are reduced.

4.1 Where to measure – the process class

The decision which processes should be enriched by checks for information quality can be made easier by classifying the processes by *process classes*. Defined process classes are:

- **supply** – forward information into the system (application)
- **acquisition** – accepts information from external suppliers and injects it into the system
- **creation** – manual data entry
- **processing** – any stage that accepts input and generates output
- **packaging** – information is collected, aggregated and summarized for reporting purposes

- **analysis** – the point where human interaction is required
- **delivery** – the point where packaged information objects are delivered to a known data consumer



Figure 4: Definition of process classes⁵

If a relevant process class can be found for a business process, a quality gate – a check-point for information quality – of a certain type should be designed into the process. For every process class a corresponding quality gate class – the appropriate “type” of the quality gate – is defined. For the current application at E.ON a 1:1-relation between process class and quality gate has proven to be suitable. The quality gate class defines a set of information quality attributes to be checked within the business process.

4.2 What to measure – the Quality Gate Class

Each quality gate class defines a set of information quality attributes (CTQs⁶) to be checked by a quality gate within a business process. Defined quality gate classes are:

- **Receiver** – receive external data or data from the supply chain
- **Creator** – data entry
- **Processor** – information from different data sources is merged or manipulated
- **High profile stages** – a processing stage that consumes a large percentage of company resources
- **Collector** – information is aggregated and prepared for reporting or for storage
- **Broadcaster** – prepares information for many consumers

The following matrix shows which information quality attributes are considered to be relevant for each of the quality gate classes:

⁵ IT processes (packaging, analysis, delivery) are treated as business processes and have to be modelled accordingly.

⁶ CTQ: Critical to Quality

	Quality Gate Class (Process Class)						
	Supply-Receiver (supply)	Acquisition-Receiver (acquisition)	Creator (creation)	Processor (processing)	Profile Stage (packaging)	Collector (analysis)	Broadcaster (delivery)
<i>1 -Believability</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>2 -Accuracy</i>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>3 -Objectivity</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
<i>4 -Reputation</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					
<i>5 -Valued-added</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>
<i>6 -Relevancy</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>7 -Timeliness</i>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>8 -Appropriate amount</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>
<i>9 -Completeness</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>10-Interpret-ability</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>
<i>11-Ease of understanding</i>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>12-Representational consistency</i>					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>13-Concise representation</i>					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>14-Accessibility</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>

	Quality Gate Class (Process Class)						
	Supply-Receiver (supply)	Acquisition-Receiver (acquisition)	Creator (creation)	Processor (processing)	Profile Stage (packaging)	Collector (analysis)	Broadcaster (delivery)
15-Access Security	<input checked="" type="checkbox"/>						<input checked="" type="checkbox"/>

Figure 5: Quality gate classes to define information quality attributes (example developed for E.ON)

4.3 Bringing where and what together – Quality Gates

A quality gate is a checkpoint for information quality in a business process. This may be a single step or a complete sub process. The measurement of information quality is done for the quality attributes as defined by the quality gate class of the quality gate. The measure itself can be an organizational assessment or just a technical monitoring task.

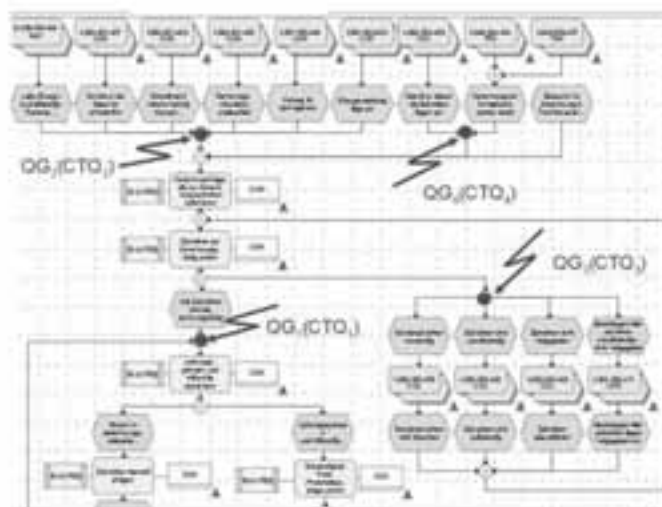


Figure 6: Quality gates in business processes

Whether information quality of a business object (resp. a group of business objects) is appropriate or not can be determined by comparing the outcome of the measure to a defined quality threshold. This threshold can be derived from the business value class of the business object and the quality gate class of the measuring quality gate.

4.4 Measure thresholds

The predefined threshold for information quality, i.e. the interpretation of what “appropriate” information quality is in a given context, is derived from the business value class of the business object under assessment and the quality gate class of the quality gate.

The threshold values are calculated by weighted quality dimension attributes (CTQs) within the quality gate classes. The basic idea: More critical information – i.e. a business object with higher business value – has higher requirements for information quality, and higher information quality requirements lead to a higher quality threshold limit (expressed in %). Figure 5 shows an example for a quality threshold matrix.

	Quality Thresholds for Business Value Classes				
Quality Gate Classes	critical information [%]	business performance information [%]	essential information [%]	sensible information [%]	non-critical information [%]
Supply Receiver	91-100	71-90	31-70	10-30	0
Acquisition Receiver	81-100	61-80	21-60	10-20	0
Creator	81-100	61-80	21-60	10-20	0
Processor	81-100	61-80	21-60	10-20	0
Stages	81-100	61-80	21-60	10-20	0
Collector	81-100	61-80	21-60	10-20	0
Broadcaster	81-100	61-80	21-60	10-20	0

Figure 7: Quality thresholds

4.5 Example

A business process acquires external market data. This process can be classified as “acquisition”-process. The corresponding quality gate class is “acquisition receiver”. Having identified the quality gate class one knows the relevant quality attributes and can measure information quality for these attributes. In order to decide whether the result is appropriate, the business value class of the business object “market data” must be considered. Let us assume, it is “business performance information”. After all this information

has been collected one can lookup the aimed threshold, which is 61% to 80% and compare it to the result.

The single steps at a glance:

- 1. identify the process class of the business process
- 2. identify the corresponding quality gate class
- 3. model and implement a quality gate in the relevant business processes
- 4. identify the business value class of the business object to be assessed
- 5. perform a quality check for the business object according to the relevant quality attributes derived from quality gate class
- 6. perform quality evaluation for the business object according to the aimed threshold derived from quality gate class and business value class

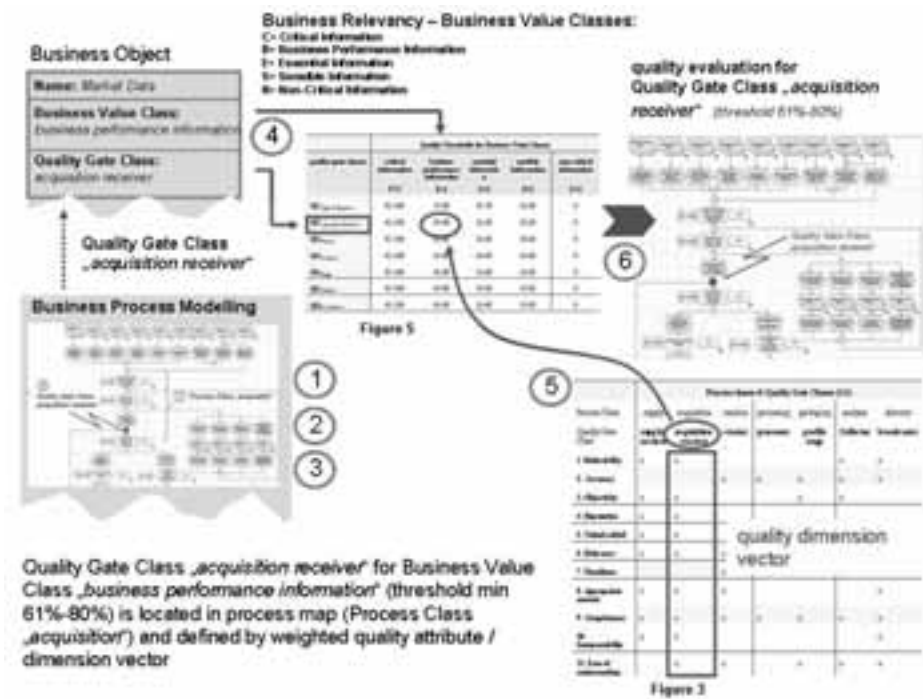


Figure 8: Quality gate measures in context with business objects and business process modeling definitions

5 Information lifecycle management

The fundamental idea behind information lifecycle management is the fact that the value and therefore the patterns of information access of most information changes over time. Information provisioning should follow this in order to fulfill business requirements

efficiently which means to align the business value of information to the most appropriate infrastructure and services. Information lifecycle managements consists of the supply of policies, processes, best practices and services to manage information in a consistent manner [Bi04].

There are many approaches to information lifecycle management provided by the storage industry, such as system managed storage [Ge99], [YCA99], storage resource management [Th03], hierarchical storage management or even information lifecycle processes [Gi03]. To meet the business requirements we choose a combined approach based on the framework of SNIA⁷ [Pe04], [Sn04] and the process descriptions of BITKOM⁸ [Bi04].

The active management of information over their whole lifecycle comprises the implementation of a rule set composed of methods, processes and technology. The business value of information faces the optimization of information service and storage costs. The basic objects of treatment are business objects defined by the information architecture.

5.1 Requirements

The most important reasons to establish information lifecycle management are the increase of information growth, storage and service costs, complexity and the regulatory constraints.

- *Increase of information growth*: The annual increase of information affects all industries. Data bases grow on average between 100% and 125% per year, while static data grow by 60%. Digitally content (films, music, etc.) grows by 50% per year [Da04], [Ly03].
- *Storage and service costs*: The cost of storage services is increasing between 23% – 30% annually, despite the fact that hardware costs are decreasing every year.
- *Complexity*: The complexity of infrastructures is increasing. New technologies like hierarchical storage management facilities, online and near line storage or enterprise information infrastructures are producing new requirements for IT management and services.
- *Regulatory constraints*: Regulations and laws require specific information management activities and their audits.

⁷ SNIA: Storage Networking Industry Association

⁸ BITKOM: Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.

5.2 Information lifecycle phases

The information lifecycle consists as minimal requirement of the four phases Operational, Analytical, Archived and Removed. Every entity of a business object used within the enterprise is at a given time in one of the four lifecycle phases. The change conditions from one lifecycle phase to another are defined by lifecycle phase rules. These rules are to be applied to business objects. As an example, safekeeping and retention time-spans are defined as lifecycle phase rules.

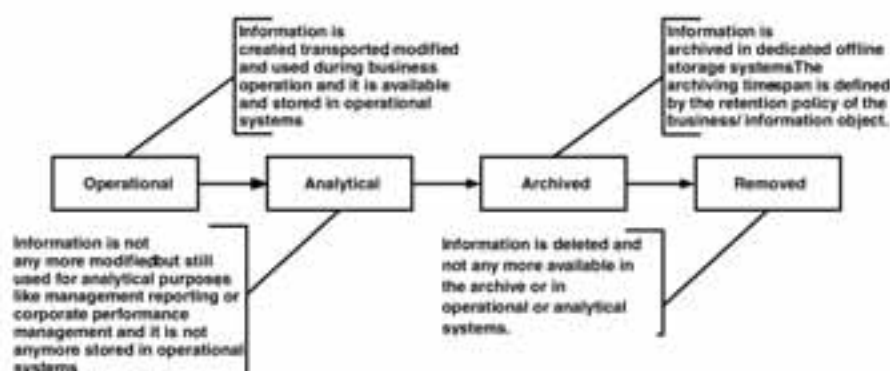


Figure 9: The information lifecycle

5.3 Information business value classes

The information business value classes are an assignment of value to business objects from a business perspective. The main reason for this classification is the possibility to treat information objects according to their relevance. This treatment is supported by constraints for infrastructure and operating service requirements deducted from the business value class. Business value classes are an integrated component of information lifecycle management as well as of information quality management.



Figure 10: The information business value classes

The objective assignment of value to information is difficult, because of the missing accounting rules for information. Different approaches to measure information’s value are possible, such as the information productivity [Di04], the information value function [Cr97], the subjective value of information [Ra03], no price for information [Ba96], or the cost of missing information [Re98]. We used a purely subjective approach. It is the responsibility of the business to assign value to a given information object.

5.4 Activities in the context of information lifecycle management

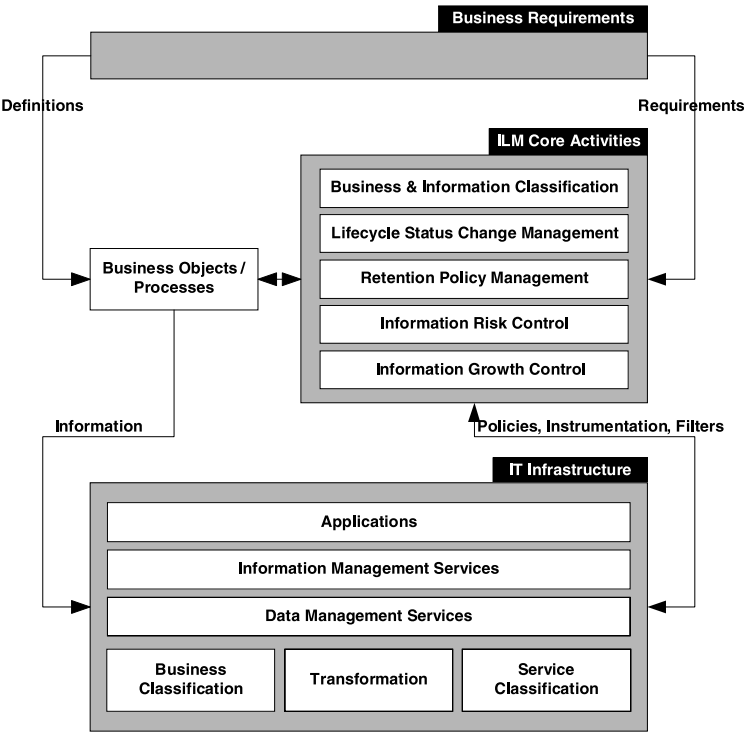


Figure 11: Information lifecycle management core activities

The information lifecycle management core activities are focused on the business objects and business processes defined by the enterprise architecture. Business requirements are the basic constraints of all core activities. The core activities are business and information classification (assigning business value classes to information objects), lifecycle status change management (applying phase rules to information objects), retention policy management (controlling regulatory rules), information risk control (comparing information business value with the actual information treatment done by IT operations), and information growth control (monitoring and forecasting).

6 Implementation at E.ON – The current state

The organizational framework and methods described above are considered as a basic framework for information management at E.ON. Due to the existing organizational structure which consists of regional or functional responsible market units and subsequent business units the implementation has to be unit-specific. Basically, an approach of minimal changes to existing organizations is pursued since it improves acceptance and minimizes reservation [Se07a], [Se07b].

Currently, one business unit which has the role of a pilot has implemented the complete framework consisting of the relevant roles and an enterprise architecture covering business process modeling, information architecture, and application architecture. The implementation for this unit has shown that remarkable efforts in time and man power are necessary to realize a complete implementation.

Because the realization of information issues, like information quality, is often time critical for future roll-outs it is planned to implement the organizational framework and a kind of ‘light’ information architecture in the context of major corporate projects. The basic idea is that during project activities the persons representing the information key roles of an information architect or information manager can be easily identified. During the project roll-out the identified persons inherit these roles for daily business and keep them alive. To minimize the initial efforts for the creation of an information architecture the relevant business object meta-information can be collected using the BO-Profile in form of a spreadsheet-table (see chapter 3). Then, independent from time-critical project requirements this information base can be transferred step by step into an information architecture tool.

7 Summary

E.ON as one of the major companies in the European utility sector identifies information as a business-critical asset. To cope with the resulting information management requirements, an information governance and a corresponding methodological framework was developed. A major element of the organizational framework is the role model which consists of three information related roles. These are an information governance board, information architects and information managers. In order to ensure the important connection to business processes and IT the additional roles of technical information stewards and business information stewards are established.

The organizational framework is accompanied by a set of generic process models for information quality and information lifecycle management and a meta-information repository. This repository, the so-called business object (BO-) profile, can be assumed as central part of an information architecture. It serves for collecting and storing BO-meta-information. In the context of information quality and information lifecycle management relevant BO-meta-information items are information business value class and lifecycle phase rules. The developed process models need only the above mentioned BO-meta-information as necessary external input. With this framework information quality initiatives or lifecycle management can be carried out with minimum additional efforts.

The developed framework consisting of the role-concept, the information architecture and the connection to business process modeling and application architecture is successfully implemented in one E.ON business unit. Basing on these experiences a future roll-out of the concept in combination with major corporate projects is planned. In order to increase acceptance a strategy of minimum impact on existing organizations is pursued.

8 References

- [Ba96] Bauwens, M.: The Three Laws of the Cyber-Economy, CMC Magazine, 1996.
- [Bi04] BITKOM: Leitfaden zum Thema „Information Lifecycle Management“, 2004.
- [Cr97] Cramer, M. L.: Measuring the value of information, infoWARcon'97, 10.-12.9.1997 Vienna, Virginia.
- [Da04] Data Mobility Group: Policy-based Data Management: the Cornerstone of ILM, Data Mobility Group White Paper, 2004.
- [Di04] Dignan, L.: Who Led the Baseline 500 and Why, Baseline Magazine October 18, 2004.
- [Ge99] Gelb, J. P.: System Managed Storage, IBM Systems Journal, Vol 28, No 1, 1989.
- [Gi03] Giesselbach, M.: Return-on-Investment durch Information Lifecycle Management, 2003 Storage Technology Corporation.
- [Lo01] Loshin, D.: Enterprise Knowledge Management. The Data Quality Approach. Morgan Kaufmann, San Diego, 2001.
- [Ly03] Lyman, P. et. al.: How Much Information?, University of California, School of Information Management and Systems, 2003.
- [Pe04] Peterson, M.: Information Lifecycle Management: a vision for the future, SNIA 29.3.2004.
- [Ra03] Rafaeli, S. R.; Daphne, R.: Experimental Investigation of the Subjective Value of Information in Trading, Journal of the Association for Information Systems Volume 4, 2003.
- [Re98] Redman, T. C.: The Impact of Poor Data Quality on the Typical Enterprise, Communications of the ACM, Feb 1998.
- [Se07a] Seiner, R. S.: The Data Stewardship Approach to Data Governance: Chapter 6. TDAN.com (<http://www.tdan.com/view-articles/5604>), 2007.

- [Se07b] Seiner, R. S.: The Data Stewardship Approach to Data Governance: Chapter 7. TDAN.com (<http://www.tdan.com/view-articles/6173>), 2007.
- [Sn04] SNIA: ILM Definition and Scope – An ILM Framework, SNIA Data Management Forum 29. July 2004.
- [Th03] Thurnhofer, R.: Storage Resource Management hilft Kosten sparen, Funkschau Nr. 20, 2003.
- [YCA99] Yamazaki, T.; Casas, J. M.; Asselin, D.: Enterprise Storage Solution, IBM International Technical Support Organisation, December 1999.

BI-Governance – Königsdisziplin strategischer Business Intelligence

Uwe Trost, Martin Zirkel

Steria Mummert Consulting AG
Business Intelligence Solutions
Domagkstraße 34
D-80807 München
{uwe.trost | martin.zirkel}@steria-mummert.de

Abstract: Die Entwicklung einer formalen BI-Strategie ist das Mittel, um BI nach den Zielen des Unternehmens auszurichten. Die Umsetzung einer BI-Strategie ist undenkbar ohne die Unterstützung einer BI-Governance. Die inhaltliche Ausgestaltung der BI-Governance als strategische Maßnahme einer BI-Strategie ist ein effizienter Weg, Governance-Strukturen im BI-Umfeld aufzubauen. Zentrale Elemente einer erfolgreichen BI-Governance-Implementierung sind transparente Entscheidungsstrukturen in den relevanten Handlungsfeldern vor dem Hintergrund einer abgestimmten Gesamtorganisation sowie ein geschlossenes Prozessmodell über alle beteiligten Ebenen von Business Intelligence. Der nachstehende Beitrag vermittelt einen Überblick über die wesentlichen Elemente einer BI-Governance.

1 Einleitung

Der vorliegende Beitrag entstand vor dem Hintergrund mehrerer konkreter Unternehmenssituationen, die in der praktischen Umsetzung ihrer BI-Strategie jeweils unterschiedliche Schwerpunkte legten, der Beitrag selbst ist jedoch unternehmensneutral verfasst. BI-Governance umfasst die Grundsätze, Verfahren und Maßnahmen, die sicherstellen, dass mit Hilfe von Business Intelligence die Informationsbedürfnisse des Unternehmens zur Erfüllung der Geschäftsziele abgedeckt werden, die Ressourcen verantwortungsvoll eingesetzt und Risiken angemessen überwacht werden.

BI-Governance hat demnach die Organisation, Steuerung und Kontrolle von Business Intelligence im Unternehmen zum Inhalt.

Zwischen einer BI-Strategie und der BI-Governance besteht nach Einschätzung der Autoren folgende ambivalente Beziehung: Eine BI-Strategie definiert die inhaltlichen Aufgaben der BI-Governance, die ihrerseits u. a. für die Prozesse einer BI-Strategie (Planung, Erstellung, Revision, Kontrolle) verantwortlich zeichnet. Strategy first, Governance follows.

Eine BI-Strategie gewährleistet über die konsequente Zielableitung aus der Unternehmensstrategie die Ausrichtung von Business Intelligence entsprechend den Zielen des Unternehmens [TrZi06]. Im Nachfolgenden wird der Begriff BI-Landschaft als Synonym verwendet, um alle im Unternehmen existierenden BI-Systeme zu adressieren. Die BI-Governance hat demnach die wichtige Aufgabe, eine strategiekonforme Ausgestaltung der BI-Landschaft sicherzustellen und die Transformation der BI-Landschaft in den Sollzustand zu koordinieren.

Die Beschreibung der Aufgaben der BI-Governance als eine strategische Maßnahme im Gesamtrahmen einer ganzheitlichen BI-Strategie ist daher ein effizienter Weg, um zu leistungsfähigen Governance-Strukturen für Business Intelligence im Unternehmen zu gelangen, wobei das vollständige Spektrum der Aufgaben erst im Gesamtkontext sichtbar wird.

Die Ausgestaltung BI-Governance ist die Planung aller Maßnahmen, die notwendig sind, um

- BI fortwährend und langfristig auf die Unternehmensziele und -prozesse strategisch auszurichten und hierbei
- die BI-Ressourcen (Mitarbeiter, Systeme und finanzielle Mittel) verantwortungsvoll und nachhaltig einzusetzen und
- die BI-Risiken zu minimieren und zu optimieren und
- eine Messung der BI-Performance zu ermöglichen.

Für einen praktikablen Umgang mit dem Begriff wird das Wesentliche einer BI-Governance in einem Leitsatz – die 5 W's der BI-Governance – zusammengefasst:

Welche Entscheidungen sind durch **Wen** in welcher Rolle bezüglich **Was** (→ BI-Objekt) **Wann** (→ BI-Life-Cycle) **Wie** (→ Prozess) zu treffen.

Thesen

These 1: *Eine BI-Governance ist, isoliert betrachtet, von allen strategischen Maßnahmen einer BI-Strategie die Wichtigste.*

Die Formulierung der These impliziert, dass die BI-Governance als Teil eines strategischen Gesamtrahmens zu sehen ist, d.h. auch für Fragestellungen der BI-Governance sind Ist-Analysen durchzuführen und die Ergebnisse zu bewerten. Weiter sind messbare Ziele zu formulieren, die in der anschließenden Design-Phase zur Ausgestaltung des Solls führen, den Maßnahmen der BI-Governance-Strategie. BI-Governance ist deswegen die zentrale Disziplin, da hier die prozesstechnischen Hebel und Entscheidungsstrukturen definiert werden, die notwendig sind, um den durch die BI-Strategie ausgelösten Veränderungsprozess verbindlich in der Organisation zu verankern.

These 2: *Es gibt keinen allgemeingültigen BI-Governance Ansatz.*

Je nach Position und Einstellung zu Fragen der Business Intelligence wird sich auch eine unterschiedliche Erwartungshaltung/Perspektive auf die BI-Governance ergeben. Aus IT-Sicht stehen u. a. Fragestellungen der Konsolidierung der BI-Landschaft oder eines BI-Betreibermodells im Vordergrund, während der Fachbereich z.B. über ein Anforderungsmanagement eine optimale Unterstützung seiner betriebswirtschaftlichen Prozesse fordert.

2 BI-Governance-Framework

Fragestellungen der BI-Governance sind von Natur aus weit reichend und je nach Schwerpunkt der BI-Strategie wird auch der Schwerpunkt der BI-Governance variieren. Für gewöhnlich dürften die handelnden Personen in der Ausgestaltung dieser strategischen Maßnahme Neuland betreten und es stellt sich die Frage, wie und wo man anfängt. Das nachstehende BI-Governance-Framework unterstützt insofern die ersten Schritte durch die Aufzeigung des gesamten Handlungsrahmens und Lieferung der Impulse für erste Handlungen im Unternehmen.

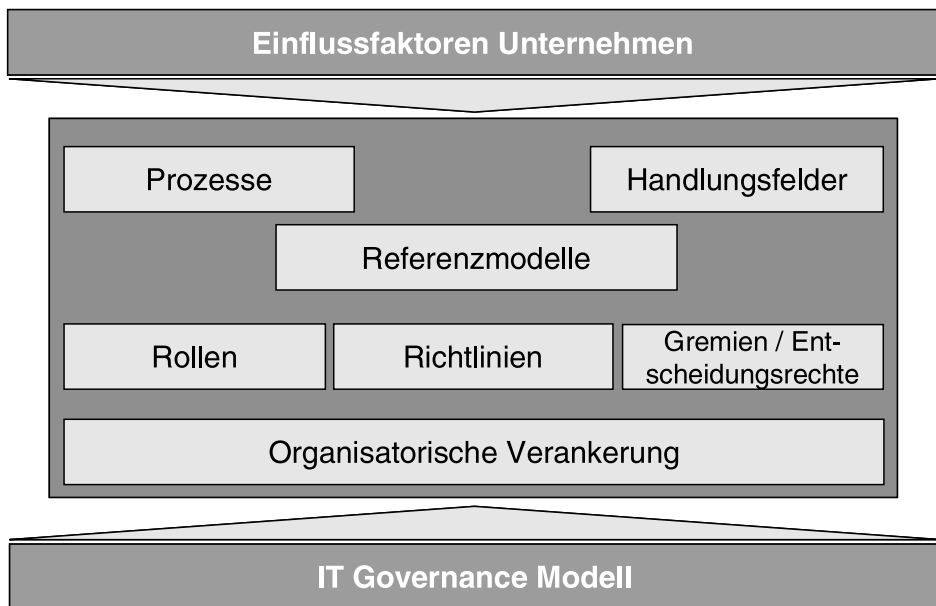


Abbildung 1: BI-Governance-Framework

Das dargestellte BI-Governance-Framework beinhaltet die wesentlichen Elemente der BI-Governance und ist seinerseits Bestandteil des BI-Strategie-Entwicklungsframeworks biSE®. Demnach ist eine BI-Governance ein organisatorisches Rahmengerüst, das im Spannungsfeld zwischen Unternehmensfaktoren und der IT-Governance existiert.

Im Folgenden wird zunächst das Rahmengerüst betrachtet. Danach werden die einzelnen Elemente des Frameworks beleuchtet.

2.1 Einflussfaktoren des Unternehmens

Einflussfaktoren des Unternehmens sind das Marktumfeld des Unternehmens und die aktuelle Unternehmensstrategie sowie die Unternehmenskultur. Beide beeinflussen nachhaltig die Ausgestaltung einer BI-Governance. Typischerweise setzt ein Unternehmen der ‚Old Economy‘ synergistisch auf Skaleneffekte und hofft dadurch Kostenvorteile zu heben. Agieren Unternehmen in Wachstumsmärkten, muss die BI-Governance sicherstellen, dass die geforderte Agilität vorhanden ist. Mergers und Aquisitions prägen moderne große Konzernstrukturen, in der Gesamtsicht lässt sich das Unternehmen häufig als ein Konglomerat unterschiedlicher Einheiten darstellen, die relativ selbständig am Markt auftreten. Hier ist das Zusammenspiel der Einheiten durch die BI-Governance zu gewährleisten. Die aktuelle Situation des Unternehmens wirkt ebenfalls klassifizierend – vom Überlebenskampf (Kostenkontrolle) hin zu den Gewinnern (Ausbau des Vorteils) werden die Governance-Modelle variieren.

Neben dem Marktumfeld steuern strategische Werttreiber und Wettbewerbsstrategien die Ausgestaltung der BI-Governance. Unter der Unternehmensstrategie ‚Operational Excellence‘ werden kostenoptimale Betriebsprozesse von der IT (BI) erwartet; überlegene Produkte oder Konzentration auf den Kunden sind Differenzierungsansätze, bei denen das Kostenargument gegenüber einer notwendigen und daher geforderten Flexibilität durch BI in den Hintergrund tritt.

Fazit: Das Unternehmen selbst und sein Marktumfeld definieren grundlegende Anforderungen an eine BI-Governance.

Auf dieser Grundlage kann man anschließend in einem bewertenden Abgleich gegen Standardstrategien festhalten, ob das vorgefundene BI-Governance-Modell tatsächlich optimal für die Business Intelligence im Unternehmen ist.

2.2 IT-Governance-Modell

Ein bestehendes IT-Governance-Modell ist, sofern vorhanden, die zentrale Referenz für die Ausgestaltung einer BI-Governance. Im Idealzustand sind dabei lediglich im Detail BI-spezifische Ableitungen zu machen. In der unternehmerischen Realität ist der Reifegrad der IT-Governance unterschiedlich weit fortgeschritten. Dort wo IT-Governance draufsteht, ist nicht immer IT-Governance drin. Die Umsetzung einer BI-Strategie und das damit verbundene Regelwerk benötigen eine handlungsfähige BI-Governance.

2.3 Organisatorische Verankerung

Die organisatorische Verankerung der BI-Governance wird in den Unternehmen unterschiedlich gehandhabt. In der Lesart einer dominanten IT-Governance wird die Zuständigkeit für das Thema BI-Governance als Spezialisierung gesehen. Vorteilhaft sind hier zweifelsohne vorhandene eingespielte Strukturen, das organisatorische Setup der BI-Governance dürfte reibungslos verlaufen. Auf der anderen Seite geht die Initiative einer BI-Strategie häufig mit dem Aufbau eines BI-CC einher. Die BI-Governance wird dann häufig eine Aufgabe innerhalb des BI-CC [So08]. Die Zentralisierung und die Nähe zu den strategischen Maßnahmen der BI-Strategie erweisen sich hier als großer Vorteil. Man muss an dieser Stelle jedoch auch auf die Grenzen zentraler Ansätze verweisen. Unternehmenskonglomerate bestehen häufig aus vormals unabhängigen Einheiten mit unterschiedlichen Steuerungsmodellen der IT. In diesem Fall tritt neben der ‚schwächeren‘ Verbindlichkeit zentraler Standards der organisatorische Overhead eines konzernweiten Modells der Zusammenarbeit als zusätzliche Herausforderung in Erscheinung.

2.4 Richtlinien

Eine wesentliche gestalterische Aufgabe der BI-Governance ist die Formulierung von verbindlichen Richtlinien und, als schwächeres Pendant, Leitlinien, die in den Prozessen entlang des BI-Life-Cycles zur Anwendung gebracht werden sollen. Beispiele für Richtlinien sind Vorgaben für den Umgang mit Information im Unternehmen oder die Ausgestaltung eines Phasenmodells über den gesamten Life-Cycle an sich.

2.5 Rollenmodelle

Rollenmodelle erfüllen eine wichtige Funktion in einer modernen Unternehmensorganisation, da Rollen mit bestimmten Aufgaben verbunden sind und in der prozesstechnischen Modellierung eingesetzt werden. In der Praxis zeigt sich, dass hier weniger meist mehr ist. Eine inflationäre Ausweitung des Rollenbegriffs ist nicht zielführend, daher sollte man bei der Identifikation von notwendigen Rollen in einem beliebigen Kontext der Business Intelligence so weit wie möglich auf bereits existierende Rollen zurückgreifen.

2.6 Gremienstruktur/Entscheidungsrechte

Die Umsetzung einer BI-Strategie geht meistens mit der Initiierung eines BI-Programms einher. Die BI-Governance wird daher in der initialen Phase eines BI-Programms damit beauftragt sein, eine Gremienstruktur aufzusetzen, in der die wesentlichen Themen der BI-Strategie bearbeitet und über eine Geschäftsordnung entschieden werden. Mit zunehmender Etablierung des BI-Programms und fortschreitender Implementierung der BI-Strategie institutionalisieren sich diese Gremien als Bestandteil der Regelorganisation.

Beispiele hierfür sind:

- Gremium zwischen Fachbereich und IT zur fachlichen Steuerung der BI-Landschaft
- Gremium zwischen IT und Betrieb zur Optimierung des Betreibermodells von BI-Systemen
- Gremien innerhalb der IT zur Bewältigung der gemeinsamen Aufgaben

Die Ausgestaltung einer Gremienstruktur ist ein Muster ohne Wert, wenn nicht eine präzise Vorstellung erarbeitet wird, wer was an welchem BI-Objekt zu sagen hat, welche Entscheidungsrechte notwendig sind. Die BI-Governance selbst trifft operativ keine Entscheidungen, hierfür ist die Führungsebene zuständig.

BI-Objekte in diesem Sinne können sein: BI-Strategie, BI-Life-Cycle, BI-Architektur, BI-Systeme, BI-Portfolio, etc. Grundsätzlich können BI-Objekte alles sein, was irgendwie „beregelt“ werden muss, damit die BI-Landschaft konform zur BI-Strategie verändert werden kann. Die Etablierung einer BI-Governance löst die bereits bestehenden Prozesse ab. Eine BI-Governance ist gut beraten, sich auf die wirklich wesentlichen Entscheidungsrechte zu beschränken, um sich nicht in unnötigen Details zu verlieren und trotzdem sicherzustellen können, dass die Prozesse unter Kontrolle sind.

2.7 Referenzmodelle

Referenzmodelle vermitteln im IT-Governance-Umfeld zwischen Aufbauorganisation und ablauffähigen Prozessmodellen. Aufgrund des generischen Grundcharakters kann keines der etablierten Referenzmodelle herangezogen werden, um sinnvolle und anwendbare BI-Governance-Prinzipien direkt abzuleiten. Gleichwohl sind alle Modelle einen vertiefenden Blick wert, denn sie vermitteln den Eindruck, worauf es in einer operativen BI-Governance ankommt: Rollen, Prozesse, Kontrollen, Regelstrukturen. Das Modell CobiT hat das Ziel, eine Kontrolle über den IT-Prozess zur Erfüllung der Businessanforderungen mittels Definition von IT-Domänen und Standard-IT-Prozessen (inkl. Sammlung von Kontrollzielen) bereitzustellen [ITGI07]. Die Ergänzung von CobiT durch ValIT stellt strategische Themen wie die Investmententscheidung und die Realisierung von Nutzen durch IT stärker in den Vordergrund. ITIL (Service-Management) ist ein de-facto-Standard im Run-Bereich. Der Großteil der Unternehmen setzt daher auch auf maßgeschneiderte Individual-Modelle [FoRe04], [FoRe05].

2.8 Handlungsfelder

Unabhängig von der persönlichen Einschätzung der Referenzmodelle ist ein wichtiger Aufgabenbestandteil, die zentralen Handlungsfelder einer BI-Governance inhaltlich abzugrenzen. In der nachstehenden Abbildung sind die Big Points der BI Governance zusammengefasst. In allen zentralen Handlungsfeldern lassen sich Aktivitäten aus den Bereichen Planung, Organisation, Kontrolle und Steuerung identifizieren. Auch kann man bei der BI-Governance nach operativen und strategischen Prozessen unterscheiden.

Im Folgenden werden die einzelnen Handlungsfelder kurz erläutert. Die Betrachtungsperspektive hierbei ist: Was muss man tun, um Business Intelligence im Unternehmen strategisch auszusteuern?

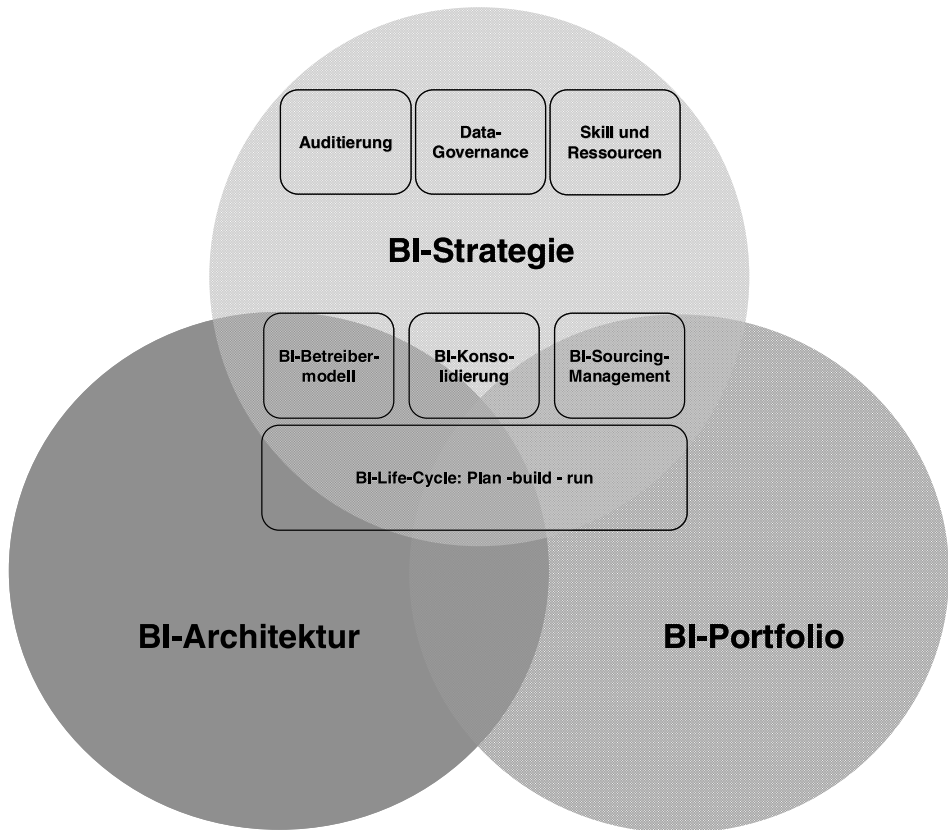


Abbildung 2: Zentrale Handlungsfelder der BI-Governance

BI-Strategie

Business Intelligence auf ein strategisches Level zu heben bedeutet, eine **BI-Strategie** zu entwickeln. Verantwortung für die BI-Strategie heißt dabei nicht nur Koordination und Erstellung der BI-Strategie, sondern auch Kontrolle der Umsetzung. Die Umsetzung einer BI-Strategie erfolgt über ein Set strategischer Maßnahmen, die häufig über ein BI-Programm gebündelt werden. In der operativen Konsequenz bedeutet dies, dass die BI-Governance über ein geeignetes Instrumentarium den Fortschrittsgrad und die Wirksamkeit der strategischen Maßnahmen kontrollieren können muss (z.B. über eine Balanced Scorecard).

BI-Portfolio

Business Intelligence auf ein strategisches Level zu heben bedeutet, sämtliche fachliche Anforderungen zu sammeln und entsprechend den Unternehmenszielen einer Priorisierung zu unterziehen, ein **BI-Portfolio** mit unterschiedlichen Dimensionen (strategische und monetäre Bewertung, Risiko- und Ressourcenbewertung) aufzubauen.

Die BI-Governance spielt auf Ebene des Anforderungsmanagements den Mittler zwischen den Welten und bringt die Informations-Anforderungsseite mit der Informations-Bereitstellungsseite zusammen und etabliert eine entsprechende Gremienstruktur.

In der Praxis umfasst das zum Einsatz kommende BI-Portfolio-Instrumentarium häufig mehrere Einzelportfolios, die zueinander in Beziehung stehen und verwendungsorientiert unterschiedlich aufgebaut sind. Mit Hilfe eines BI-System-Portfolios werden bestehende BI-Systeme beurteilt. Bei entsprechender Wahl der Dimensionen (Achsen) eines Systemportfolios können auch Ist- und Soll-Positionen für einzelne BI-Systeme identifiziert werden. Die zu identifizierenden Maßnahmen führen ultimativ zu Projekten oder Projektideen, die in einem BI-Projekt-Portfolio zueinander und zur Umsetzung neuer fachlicher Anforderungen betreffenden Projekten in Beziehung zu setzen sind.

Nur unter Betrachtung des gesamten BI-Portfolios lässt sich aufzeigen, wo aus strategischer Sicht Lücken bestehen und wie viel investiert werden muss und welche Maßnahmen unter Berücksichtigung aller Dimensionen tatsächlich lohnend sind. Das BI-Portfolio ist daher ein strategisches Gestaltungsinstrument zur zielgerichteten und strategiekonformen Ausgestaltung der BI-Landschaft.

BI-Architektur

Business Intelligence auf ein strategisches Level zu heben bedeutet, einen erweiterten Begriff von Architektur einzuführen. BI-Architekturmanagement heißt BI-Perspektiven einer Unternehmensarchitektur aufzubauen [Ni05]. Die BI-Perspektiven ergeben eine strukturierte Sammlung von Plänen für die Gestaltung der BI-Landschaft – ausgehend von Zielen und Strategien auf Ebene der Geschäftsarchitektur über Fach- und Anwendungssichten zu einer technischen Systemarchitektur. So ist die Etablierung technischer Referenzarchitekturen im BI-Umfeld ein Standardisierungsinstrument der BI-Governance.

Auf der gestalterischen Ebene ist ein BI-Bebauungsplan ein Element des BI-Architekturmanagements und damit der BI-Governance. Unter der Annahme, dass eine BI-Landschaft durch Komplexität und Heterogenität geprägt ist, schlägt sich eine gewisse Bandbreite an vorhandenen Strategien, Prozessen, Architekturen und Technologien nieder in einem erhöhten Aufwand bei der Ausplanung konzernweiter BI-Initiativen sowie der Gefahr von Redundanzen, Überschneidungen und letztlich Fehlinvestitionen in BI.

Aus der BI-Strategie und deren Teilstrategien wird eine Soll-Bebauung abgeleitet. Der Weg vom jeweiligen Ist zum Soll wird durch eine Roadmap geeigneter Maßnahmen repräsentiert. Diese Maßnahmen fließen gemeinsam mit den aus neuen fachlichen Anforderungen resultierenden Vorhaben in eine Portfolio-Betrachtung ein, anhand derer die durchzuführenden Projekte identifiziert, priorisiert und kombiniert werden. Hier wird deutlich wie eng BI-Strategie, BI-Architektur und BI-Portfolio als Instrumente der BI-Governance zusammenhängen.

Auditierung

Business Intelligence auf ein strategisches Level zu heben bedeutet, Transparenz über die BI-Landschaft zu haben. Wenn strategische Maßnahmen Veränderungen in der BI-Landschaft erzeugen, dann ist es auch Aufgabe der BI-Governance dafür zu sorgen, dass die konkreten Auswirkungen (z.B. Reduzierung der Betriebskosten um x %) informationstechnisch erhoben und historisiert abgespeichert werden. Erst wenn Transparenz über den aktuellen Zustand der BI-Landschaft gegeben ist, kann davon gesprochen werden, dass die BI-Landschaft beherrsch- und damit steuerbar ist. Eines der wirksamsten Instrumente einer BI-Governance ist daher ein BI-System, in welchem alle Informationen im Kontext Business Intelligence gebündelt, aktuell vorgehalten und ausgewertet werden, BI für BI.

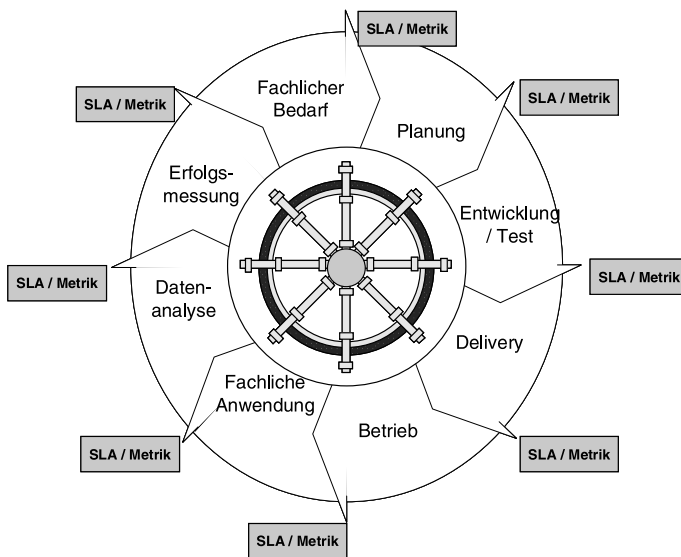


Abbildung 3: BI-Life-Cycle-Prozess

BI-Life-Cycle

Business Intelligence auf ein strategisches Level zu heben bedeutet, die bloße Vorstellung von Business Intelligence als eine Art von Software-System zu erweitern durch die Vorstellung von Business Intelligence als ein kontinuierlicher (Geschäfts-) Prozess.

In der Ausgestaltung der BI-Governance wird die Entscheidungsfindung in sämtlichen Entscheidungsbereichen im Rahmen des BI-Life-Cycles eines BI-Systems mit prozesstechnischen Hebeln ausgestaltet. Die prozessuale Sichtweise erlaubt es sämtliche Prozessschritte – beginnend beim fachlichen Bedarf über die Entwicklung eines BI-Systems hin zu einer nachgelagerten Erfolgsmessung – mit SLAs auszugestalten und so z.B. die Verwendung von Softwarestandards für die Entwicklung verbindlich vorzuschreiben und dies über Metriken zu kontrollieren.

Data Governance

Business Intelligence auf ein strategisches Level zu heben bedeutet, sich des Wertes guter Information bewusst zu sein. Informationen und Daten sind ein zentraler Wirtschaftsfaktor mit zunehmender Bedeutung und müssen unterschiedlichsten Ansprüchen genügen, beginnend bei konsistenten Begriffsdefinitionen hin zu regulatorischen Anforderungen im Sinne der Compliance. Die Relevanz einer Data Governance wird im Rahmen eines Assessments schnell offensichtlich.

Auf Ebene der Data Governance kann das BI-Governance-Framework nahezu vollständig zur Anwendung gebracht werden. Data Governance umfasst dabei den kompletten Life-Cycle der Informationen, von der Erst-Anlage bis zum Löschen der Daten. Ein wichtiges Element einer Data Governance ist die Definition der Rolle des Data Stewards.

Wichtige verantwortliche Aufgaben des Data Steward beinhalten:

- Sicherstellen von Datenkonsistenz
- Gewährleistung Datensicherheit
- Datenzugreifbarkeit
- Datenverfügbarkeit
- Datenqualität
- Datenauditierbarkeit
- Datenverwendungsnachweis
- Gültigkeit und Eindeutigkeit von Datendefinitionen
- Daten-Change-Management

Organisatorisch sollte der Data Steward so nah wie möglich am Entstehungspunkt der Daten sitzen, und spezifische Ziele bezüglich einer zu steigernden Datenqualität haben. Data Stewardship ist dabei ungleich Data Ownership. Der Data Steward wird unterstützt durch ausgestaltete Schnittstellen zur BI-(IT)-Governance, einer funktionierenden Gremienstruktur mit Fokus Datenqualität (Data Governance Boards), Standards, Richtlinien und Prozesse, Daten-Qualitätskennzahlen sowie Tools und Technologien.

Konsolidierung

Business Intelligence auf ein strategisches Level zu heben bedeutet, den Prozess der Konsolidierung der BI-Landschaft zu steuern. Ausgehend vom durchschnittlichen Reifegrad der Business Intelligence in den deutschen Unternehmen sind in der Vergangenheit zahlreiche projektgetriebene, fachlich isolierte BI-Systeme entstanden. Die Konsolidierung der BI-Landschaft unter Berücksichtigung der Anforderungen des strategischen Sollzustandes einer BI-Plattform ist ein komplexes Unterfangen, das über die BI-Governance gesteuert wird. Die Komplexität wird zum einen bestimmt durch die Anzahl der BI-Systeme an sich, zum anderen durch die zahlreichen Ebenen der Konsolidierung, die durch den Aufbau eines BI-Systems vorgegeben sind. Grundsätzlich lassen sich hier technische und fachliche Treiber unterscheiden. Aufgabe der BI-Governance im Kontext der Konsolidierung ist es zunächst, alle notwendigen Informationen strukturiert zu erheben (s. Auditierung), diese einem auf BI-Belange adaptierten Bewertungsschema zuzuführen und in Abhängigkeit von Dimensionen wie Risiko/Kosten eine Entscheidung für ein bestimmtes Konsolidierungsszenario vorzubereiten, sprich einen Konsolidierungsprozess zu etablieren.

Skill- und Ressourcenmanagement

Business Intelligence auf ein strategisches Level zu heben bedeutet, das im Unternehmen vorhandene Know-How über BI zu steigern. Die BI-Governance ist im Sinne eines Skill- und Ressourcenmanagements dafür verantwortlich handlungsfähige Strukturen aufzubauen, so dass die Skill-Level entsprechend der Benutzeranforderungen aufgebaut werden und strategische Projekte nicht an der Ressourcenfrage scheitern.

BI-Betreibermodell

Business Intelligence auf ein strategisches Level zu heben bedeutet, auf der Betriebsseite für kostenoptimale Effizienz und tragfähige Kompromisse im Bereich der Flexibilität/Qualität voranzutreiben. Die BI-Governance ist entsprechend des BI-Life-Cycle-Ansatzes dafür verantwortlich ein sauberes Delivery-Management aufzusetzen und als Grundlage für das Betreibermodell zu etablieren. An diesem Übergang von build zu run muss überprüft werden, ob die vorab definierten verbindlichen Standardisierungen (Produkt, Architektur, Softwarestandards) tatsächlich zum Einsatz kommen. Kostenoptimale Effizienz ist in erster Line das Ergebnis intelligenter Standards im BI-Umfeld. Über entsprechende Gremienstrukturen muss es die BI-Governance ermöglichen, dass Betriebsanforderungen bereits in Planungsphasen berücksichtigt und in den Build-Phasen verwendet werden.

Nachdem sich ITIL v.a. im run-Bereich als de-facto-Standard etabliert hat, besteht die Aufgabe der Entwicklung eines BI-Betreibermodells darin, die in ITIL vorgesehenen (Manager-)Rollen auf BI-Spezifika zu adaptieren und über ein ablauffähiges Prozessmodell zu verbinden. Neben Kostenoptimierung sind weitere Ziele eines BI-Betreibermodells eine Steigerung der Support-Qualität und die Reduktion von Risiken in der BI-Landschaft zu erreichen und eine Messbarkeit von Qualität/Stabilität der technischen BI-Landschaft zu etablieren. Hier sorgt die BI-Governance im Sinne einer Ent-

scheidungsvorbereitung dafür, dass der Optimierungsbedarf der BI-Landschaft aus Betriebssicht transparent wird.

BI-Sourcing-Management

Business Intelligence auf ein strategisches Level zu heben bedeutet, sich um Zukunftsaufgaben Gedanken zu machen. Outsourcing ist ein solches Thema, dem sich jeder BI-Programm-Manager aufgrund des Kostendruckes in Zukunft stellen muss [Tr08]. Neben der Begleitung der strategischen Entscheidung für ein bestimmtes BI-Sourcing-Modell wird nach Wahl einer bestimmten Option, die BI-Governance dafür verantwortlich sein, die Voraussetzungen für das Modell zu schaffen. Die Konsequenzen können dabei sehr weitreichend sein. Daten werden umso wertvoller (sensibler) für ein Unternehmen, desto intelligenter sie über Auswertungen in einen für das Unternehmen relevanten Kontext gestellt werden können. Es ist daher wahrscheinlicher, dass man für den Betrieb eines Data Warehouses einen Outsourcing-Business-Case findet, als für den Aufbau und die Bereitstellung von sensiblen Informationen durch analytische Services. Dies hat weitere Konsequenzen für eine Standard-BI-Architektur, die eine anwendungsneutrale Schicht strikt von der Auswerteschicht trennt. Es ist dann wiederum die BI-Governance, die dafür zu sorgen hat, dass die Systementwicklung architekturkonform verläuft.

2.9 Prozesse

Das BI-Prozessmodell ist ein innerhalb dieser abgestimmten Organisation wirkendes in sich konsistentes Prozessnetzwerk, welches unter anderem gewährleistet, dass die zentralen BI-Governance-Objekte (BI-Portfolio, BI-Bebauungsplan, BI-Life-Cycle, etc.) aufeinander abgestimmt sind. Die BI-Governance-Prozesse stellen die steuernden Prozesse dar und lösen die bereits bestehenden Prozesse ab.

Die BI-Strategie bildet die Grundlage für das Handeln im BI-Governance-Prozess-Netzwerk. Insbesondere gibt sie die Orientierung für die Prozesse der BI-Data-Governance, der fachlichen Planung, des BI-Portfolio-Managements sowie der ganzheitlichen und strategiekonformen Steuerung der Konsolidierung vor. Sowohl die fachliche Planung als auch das BI-Portfolio-Management und die Konsolidierung benötigen Informationen über BI-Systeme und laufende BI-Projekte.

Zu diesem Zweck wird das BI-Audit-Management mit der Informationserhebung beauftragt. BI-Data Standards (als Oberbegriff für die Ergebnistypen der BI-Data-Governance-Prozesse) sowie die BI-Referenzarchitektur werden über das BI-Lifecycle-Management an die BI-Projekte herangeführt. Der fachliche Bebauungsplanung sowie das BI-Portfolio geben im Zusammenspiel den Gestaltungsraum für BI-Projekte vor und steuern insbesondere die BI-Landschaft. Informationen über BI-Systeme und Projekte werden nicht nur im Rahmen der Audits für die zuvor angesprochenen Gestaltungsprozesse gesammelt sondern auch vom BI-Finanzmanagement und für das BI-Lifecycle-Management benötigt. Über den BI-Balanced-Scorecard-Prozess wird das komplette Prozessnetzwerk kontrolliert und die so gewonnenen Erkenntnisse fließen in den Strategieprozess ein. Das gesamte Prozessnetzwerk unterliegt dabei einem kontinuierlichen Verbesserungsprozess.

Das BI-Prozessmodell hat demnach im Sinne der BI-Strategie einen organisationsübergreifenden Charakter. Es greift BI-relevante Prozesse im Fachbereich auf, stellt die Prozesse dar, die in der IT verantwortet werden und der Umsetzung der BI-Strategie dienen und umfasst die BI-Prozesse im BI-CC, im BI-AMS und bei IT-Services.

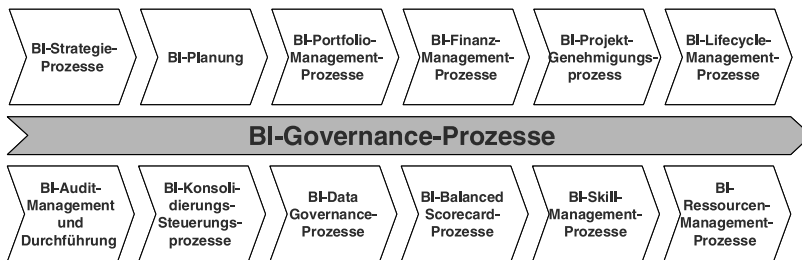


Abbildung 4: BI-Governance-Prozesse

Abbildung 4 zeigt die wesentlichen Prozessgruppen der BI-Governance. Diese decken alle bereits zuvor beschriebenen Handlungsfelder der BI-Governance ab.

3 Fazit

Zur Königsdisziplin der Business Intelligence wird die BI-Governance dann, wenn es gelingt, die in den vorab dargestellten Handlungsfeldern entwickelten **Prozesse** und Verfahrensabläufe in ein organisatorisch ‚scharf geschaltetes‘ Prozessmodell zu gießen, das alle Ebenen der Business Intelligence miteinander verbindet. Prozesse sind das Element des BI-Governance-Frameworks an dem sichtbar wird, was die BI-Governance leistet. Das Prozessmodell stellt die Einzelaktivitäten in einen Gesamtkontext und beantwortet verbindlich die Frage: Welche Entscheidungen sind durch Wen in welcher Rolle bezüglich Was (BI-Objekt) Wann (BI-Life-Cycle) zu treffen.

In der vorab dargestellten Sicht auf die BI-Governance konnten die zahlreichen Facetten des Themas nur oberflächlich gestreift werden, das BI-Governance-Framework stellt einen aus der Praxis heraus entwickelten Ordnungsrahmen dar, der einen dabei unterstützt, die Übersicht zu behalten.

Die Etablierung einer BI-Governance ist ein weitreichendes Unterfangen, das mit einer Reihe von Risiken verknüpft ist.

- Unzureichende Fokussierung auf ‚brennende‘ Themen
- Die hohe Komplexität der BI-Governance-Instrumente erfordert breites Erfahrungswissen, das selten vorhanden ist.
- Entscheidungsrechte und vor allen Dingen -pflichten werden in den Organisationen nicht immer gelebt.
- Eine schwache IT-Governance schwächt auch eine BI-Governance, schwächt eine BI-Strategie.

- Die reine Lehre dominiert pragmatischen Realismus.

In der Praxis haben sich folgende Erfolgsfaktoren herauskristallisiert.

- Verankerung der Ausgestaltung der BI-Governance als Teilstrategie einer BI-Strategie
- Aufbau einer abgestimmten Gesamtorganisation ist essentiell.
- Etablierung eines Sets aufeinander abgestimmter BI-Governance-Prozesse und -Instrumente
- Integration der Fachbereichszielsetzungen in die Ausgestaltung der BI-Governance-Instrumente
- Transparente Entscheidungsrechte und -pflichten, mit denen eine hohe Verbindlichkeit einhergeht
- Klare Zielvision bezüglich des weiteren Ausbaus der BI-Systeme
- Transparenz über die vorhandenen BI-Systeme
- Etablierung von Richt- und Leitlinien der BI-Governance in der Unternehmenspraxis
- Kommunikation der BI-Governance und deren Ergebnisse

4 Literaturverzeichnis

- [FoRe04] Forrester Research: Analyzing Europe's IT Governance – Business Technographics, Europe, Cambridge/MA, USA, 2004.
- [FoRe05] Forrester Research: Survey of 48 IT decision makers in Europe, 2005a.
- [ITGI07] IT Governance Institute: CobiT 4.1, 2007.
- [NI05] Niemann, Klaus D.: Von der Unternehmensarchitektur zur IT Governance, 2005.
- [Rü06] Rüter, A.: IT-Governance in der Praxis, 2006.
- [So08] Thorsten Sommer, Birgit Wendlandt, Uwe Trost, Martin Zirkel: Business Intelligence-Strategie bei der Volkswagen AG In (Dinter, B, Winter R., (Hrsg.)): Integrierte Informationslogistik. Springer-Verlag Berlin Heidelberg, 2008; S: 261 – 283.
- [Tr08] Ante Trbara: „Outsourcing – Strategien im Data Warehousing Umfeld am Beispiel der Deutsche Post AG“, Vortrag St. Galler Anwenderforum 18. Februar 2008.
- [TrZi06] Uwe Trost, Dr. Martin Zirkel, „Wege aus dem Informationschaos“, BI-Spektrum 2006, S. 16-19.
- [WeRo04] Weill P., Ross J.: IT Governance (2004).
- [WeRo05] Weill P., Ross J.: A Matrixed Approach to Designing IT Governance IT. In: MIT Sloan Management Review 46 (2005).

Das IT Performance Measurement Maturity Model – Ein Reifegradmodell für die Business Intelligence- Unterstützung des IT-Managements

Jörg Becker*, Ralf Knackstedt*, Jens Pöppelbuß*, Lars Schwarze**

*European Research Center for Information Systems

Leonardo-Campus 3

48149 Münster

{joerg.becker | ralf.knackstedt | jens.poeppelbuss} @ercis.de

**Deloitte Consulting GmbH

Franklinstraße 46-48

60486 Frankfurt am Main

lschwarze@deloitte.de

Abstract: Das IT-Management benötigt für die erfolgreiche Erfüllung seiner Führungsaufgaben Transparenz über die Effektivität und Effizienz des IT-Einsatzes. Ziel des IT Performance Measurement ist es, mit Hilfe von Kennzahlen und Kennzahlensystemen die IT ganzheitlich abzubilden und dabei sowohl die traditionell berücksichtigte Kostenseite als auch die Leistungsseite in die Betrachtung einzuschließen. Business Intelligence-Systeme dienen der informationstechnischen Unterstützung des IT Performance Measurement, indem sie aus verschiedensten Quellen interne Kosten- und Leistungsdaten sowie externe Marktdaten sammeln, um diese zu relevanten Informationen über die IT Performance aufzubereiten. In diesem Beitrag wird ein IT Performance Measurement Maturity Model (ITPM³) vorgestellt, das Entwicklungsstufen der Unterstützung des IT-Managements durch Business Intelligence-Systeme unterscheidet. Das Reifegradmodell bildet eine Grundlage für die Durchführung von Auditing- und die Ableitung von Weiterentwicklungsmaßnahmen in Unternehmen. Es wurde in einem iterativen Prozess und unter empirischen Überprüfungen auf Basis von Expertenbefragungen entwickelt.

1 Einsatz von BI-Systemen für ein fortgeschrittenes IT Performance Measurement

Der Einsatz von Informationstechnologie (IT) ist für viele Unternehmen ein wettbewerbsentscheidender Faktor und seine Bedeutung nimmt stetig zu. Einerseits ist die Unterstützung bestehender Geschäftsprozesse durch IT nahezu unverzichtbar geworden [ZB04, LWS06]. Andererseits bieten innovative IT-Systeme Chancen zur Steigerung der Wettbewerbsposition [HV93]. Um die IT effektiv und effizient zu gestalten, benötigt das IT-Management entscheidungsrelevante Informationen. Das klassische IT-Controlling greift an dieser Stelle zu kurz, wenn es sich nur auf die Erfassung und Verrechnung von

IT-Kosten konzentriert [MTS06]. Weiter geht der Ansatz des *IT Performance Measurement*. Dieser versucht, mit Hilfe von Kennzahlen und Kennzahlensystemen die IT ganzheitlich abzubilden, indem neben der traditionell stark beachteten Kostenseite auch die Leistungsseite, also die Performance der IT, in die Betrachtung eingeschlossen wird [GI02, BGK04, EHH06]. Konsolidierte technische Daten werden hierbei mit wirtschaftlichen Kennzahlen verknüpft [Ge02]. Des Weiteren wird durch den zunehmenden Einbezug strategischer Aspekte einer Konzentration auf die operative IT-Betriebsebene entgegengewirkt, so dass eine verbesserte Abstimmung der IT-Strategie mit der Unternehmensstrategie gewährleistet werden kann [MTS06].

Ein wirkungsvolles IT Performance Measurement verfolgt und überwacht die Umsetzung der an der Unternehmensstrategie ausgerichteten IT-Strategie, die Umsetzung von Projekten, die Verwendung von Ressourcen, die Prozess-Performance und die IT-Leistungserbringung [ITG05]. Um IT Performance Measurement erfolgreich umzusetzen, ist neben einem inhaltlich ausgereiften Konzept der Einsatz von IT-basierten Werkzeugen unerlässlich. Hierfür eignen sich insbesondere Werkzeuge aus der Kategorie der Business Intelligence-Systeme (BI-Systeme) [Ra04]. Diese sind in der Lage, aus verschiedensten Quellen interne Kosten- und Leistungsdaten sowie externe Marktdaten zu sammeln, diese Daten zu relevanten Informationen über die IT Performance aufzubereiten und durch die Bereitstellung dieser Informationen das IT-Management in seinen Führungsaufgaben zu unterstützen. Bisher spielte das IT-Management, im Vergleich zu anderen Unternehmensfunktionen in Bezug auf die Unterstützung durch BI-Systeme jedoch nur eine untergeordnete Rolle [CG04].

Um dem IT-Management ein Werkzeug zur strukturierten Fortentwicklung des BI-Systemeinsatzes an die Hand zu geben, wird im vorliegenden Beitrag mit dem IT Performance Measurement Maturity Model (ITPM³) ein Modell entwickelt, das Reifegrade in der Nutzung von BI-Systemen zur Umsetzung des IT Performance Measurement identifiziert und beschreibt. Das entwickelte Reifegradmodell schließt eine Lücke im Spektrum bestehender Reifegradmodelle, die bisher schwerpunktmäßig entweder das IT-Management oder den BI-Systemeinsatz adressieren (vgl. Abschnitt 2). Der Entwicklungsprozess des ITPM³ gliederte sich in drei Phasen, in denen es kontinuierlich verbessert wurde. Zunächst wurde eine erste Version des Reifegradmodells basierend auf existierenden Reifegradmodellen hergeleitet. Diese Version wurde in der zweiten Phase im Rahmen einer explorativen Analyse überprüft und weiterentwickelt. In der dritten Phase fand eine erneute Überprüfung im Rahmen einer Expertenbefragung durch IT-Berater der *Deloitte Consulting GmbH* statt, aus der abschließende Modifikationen resultierten (vgl. Abschnitt 3). Mit dem ITPM³ liegt erstmalig ein Reifegradmodell vor, das sich dediziert der BI-Systemunterstützung des IT Performance Measurement widmet. Seine fortgesetzte empirische Überprüfung und seine Anwendung in der Unternehmenspraxis werden Gegenstand weiterführender Forschungsarbeiten sein (vgl. Abschnitt 4).

2 Verwandte Arbeiten

2.1 Überblick

Ein *Reifegrad*, der einem Objekt zugeordnet wird, definiert die erreichte Rangstufe im Entwicklungsprozess des Objektes. Ein *Reifegradmodell* umfasst eine Folge dieser Reifegrade für eine Klasse von Objekten. Ein Reifegradmodell beschreibt somit einen antizipierten, gewünschten oder typischen Entwicklungspfad für Objekte einer Klasse in aufeinander folgenden diskreten Rangstufen, beginnend in einem Anfangsstadium bis hin zur vollkommenen Reife. Das Fortschreiten auf diesem Entwicklungspfad bedeutet i. d. R. eine stete Steigerung der Leistungsfähigkeit bzw. Qualität des betrachteten Objekts, wobei das Reifegradmodell als Skala zur Beurteilung dient. Ein Reifegrad ist durch festgelegte Merkmale des zu untersuchenden Objekts und durch die jeweils zur Erreichung des Reifegrads erforderlichen Merkmalsausprägungen definiert. Um eine nicht zu einseitige Beurteilung der Reife des betrachteten Objekts vorzunehmen, werden zumeist mehrere Merkmale bzw. Gruppen von Merkmalen in ein Reifegradmodell gefasst. Auf diese Weise erfolgt eine multidimensionale Analyse, bei der für das Objekt festgestellte Reifegrade von Dimension zu Dimension variieren können. Als Ergebnis ergibt sich so ein differenziertes Reifegrad-Profil des Objekts.

Die Anwendung des Reifegradmodells zur Ermittlung individueller Reifegrade von Objekten erfolgt i. d. R. mittels vorgegebener Richtlinien und Standards (z. B. Beschreibungen der Reifegrade, Fragebögen oder Checklisten sowie Regeln zu deren Anwendung). Zu einem gegebenen Zeitpunkt werden Beobachtungen gesammelt und validiert, um eine Zustandsaufnahme des betrachteten Objekts zu erhalten. Häufig werden Prozesse durch Reifegradmodelle beurteilt. Als Ergebnis werden dann einheitliche und überprüfbare Aussagen zu ihrem Status und zur Qualität ihrer Durchführung erwartet [Th04]. Ausgehend von der ermittelten Ist-Situation lassen sich dann Verbesserungsvorschläge und Handlungsempfehlungen ableiten [ITG05]. Der Zielzustand sollte jedoch zuvor individuell festgelegt werden. Der höchste Reifegrad kann zwar grundsätzlich als erstrebenswertes Ziel aufgefasst werden, jedoch ist eine Fortentwicklung entlang des Entwicklungspfads häufig mit beachtlichen und zunehmenden Aufwendungen (z. B. finanzieller Art) verbunden. Diese können möglicherweise im Anschluss nicht durch den Nutzen eines höheren Reifegrads kompensiert werden.

Im Kontext von Informationssystemen werden Reifegradmodelle zur Beschreibung von Lebenszyklen und zur Qualitätsbeurteilung genutzt. Ein klassisches Reifegradmodell ist die *Stage Theory* von [No79]. Dieses Modell identifiziert basierend auf empirischen Untersuchungen sechs Entwicklungsstufen der Informationsverarbeitung. Als das bekannteste Reifegradmodell kann das *Capability Maturity Model* (CMM) gelten, das 1986 bei einer Initiative des US-Verteidigungsministeriums am Software Engineering Institute (SEI) der Carnegie Mellon University in Pittsburgh entwickelt wurde. Es beschreibt Reifegrade von Softwareentwicklungsprozessen und diente für viele weitere Reifegradmodelle als Entwicklungsgrundlage. Im Folgenden werden neben dem CMM weitere ausgewählte Reifegradmodelle vorgestellt, die bereits die einzelnen Bereiche IT-Management sowie Business Intelligence getrennt betrachten. Das ITPM³ wurde

entwickelt, um diese Bereiche in einem Reifegradmodell zu fusionieren, so dass IT-Manager ein geeignetes Werkzeug zur sowohl konzeptionellen als auch technischen Fortentwicklung des IT Performance Measurement in ihrer jeweiligen Organisation erhalten.

2.2 Reifegradmodelle zum IT-Management

Zu den Reifegradmodellen, die sich im weiteren Sinne auf IT-Management beziehen, zählen bspw. das *Capability Maturity Model* (CMM) und seine Weiterentwicklung *Capability Maturity Model Integration* (CMMI) [Kn06], die Reifegradmodelle für die Prozesse im Framework *Control Objectives for Information and Related Technology* (Co-bit) [ITG05], das *IT Balanced Scorecard Maturity Model* [GS01] sowie das *IS/ICT Management Capability Maturity Framework* [Re04].

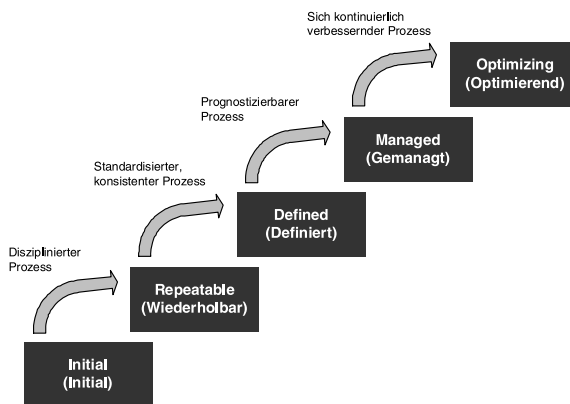


Abbildung 1: Die fünf Stufen des Capability Maturity Model [Pa93]

Das CMM unterscheidet die fünf Stufen *Initial*, *Repeatable*, *Defined*, *Managed* und *Optimizing* zur Beschreibung des Reifegrades von Software-Entwicklungsprozessen (vgl. Abbildung 1) [Pa93]. Jede Stufe beinhaltet eigene Prozessziele, deren Erreichen zu einer zunehmenden Stabilisierung des Software-Entwicklungsprozesses führt. Das übergeordnete Ziel ist die Entwicklung hin zu einer kontinuierlichen Prozessverbesserung. Aus dem CMM lassen sich Handlungsempfehlungen für eine entsprechende stufenweise Verbesserungen ableiten. Um die nächst höhere Stufe zu erreichen, müssen zunächst die Anforderungen der aktuellen Stufe erfüllt sein.

Da sich der Ansatz des CMM bewährt hatte, waren in den 1990er Jahren neben dem CMM für Software weitere Capability Maturity Models entwickelt worden [Kn06]. Diese wiesen jedoch unterschiedliche Strukturen auf, so dass ein gemeinsamer Einsatz auch dort schwierig war, wo er inhaltlich sinnvoll gewesen wäre. Daraufhin wurde das neue Projekt *Capability Maturity Model Integration* (CMMI) gestartet, welches die verschiedenen Modelle integrieren sollte. Die aktuelle Version 1.2 des CMMI wurde im August 2006 veröffentlicht [CM06].

Das Konzept des CMM wird u. a. durch das IT-Governance-Framework *CobiT* aufgegriffen, welches von der ISACA (Information Systems Audit and Control Association) entwickelt wurde und stetig aktualisiert wird [ITG05]. Es definiert 34 IT-Prozesse in den vier Bereichen *Planung und Organisation*, *Beschaffung und Implementierung*, *Betrieb und Unterstützung* sowie *Überwachung und Bewertung*. Für jeden dieser IT-Prozesse werden in Anlehnung an das CMM Reifegrade in sechs Abstufungen von *nicht-existent* (0) bis *optimiert* (5) beschrieben. Bspw. umfasst der CobiT-Prozesses *Monitor and Evaluate IT Performance* (ME1) die Festlegung von relevanten Performance-Indikatoren, eine systematische und zeitnahe Berichterstattung der Performance sowie ein promptes Handeln im Fall von Abweichungen.

Das *IT Balanced Scorecard Maturity Model* lehnt sich ebenfalls an das CMM an und dient dazu, in Unternehmen eingesetzte IT Balanced Scorecards (IT-BSC) zu beurteilen. Das Modell umfasst fünf zu durchlaufende Entwicklungsstufen, die von den ersten Schritten zur Einführung einer IT-BSC bis zur vollständigen Integration von Unternehmens-BSC und IT-BSC reichen.

Ein weiteres Reifegradmodell zum IT-Management präsentiert [Re04] mit dem *IS/ICT Management Capability Maturity Framework*. Es ist das Ergebnis eines Forschungsprojekts mit dem Ziel, ein möglichst umfassendes Modell zur Bestimmung von Reifegraden der Leistungsfähigkeit des IT-Managements zu entwickeln. Teilaspekte des IT-Managements, die bisher einzeln betrachtet wurden, sollten integriert und deren Interdependenzen berücksichtigt werden. Zur Beschreibung der Reife des IT-Managements wurden sieben Indikatoren (*Applications*, *Business-IT Relationship*, *Strategy Alignment*, *User Profile*, *Managerial Paradigm*, *Governance* und *Organization*) in das Modell aufgenommen, die jeweils drei bis fünf Entwicklungsstufen umfassen.

2.3 Reifegradmodelle zu BI-Systemen

Die zuvor präsentierten Reifegradmodelle betrachten zumeist nicht die computerbasierte Unterstützung des IT-Managements, wie sie bspw. durch BI-Systeme möglich wäre. Für letztere gibt es wiederum eigene Reifegradmodelle, z. B. die beiden *Business Intelligence Maturity Models* von [CG04] und [Ec06]. Ein weiteres relevantes Reifegradmodell ist das *Maturity Model for Performance Measurement Systems* von [WK02].

[CG04] identifizieren in ihrem *Business Intelligence Maturity Model* (biMM) die drei Schlüsselbereiche *Fachlichkeit*, *Technik* und *Organisation*, die zur Bewertung von BI-Systemen herangezogen werden können (vgl. Abbildung 2). Der Bereich *Fachlichkeit* steht für die betriebswirtschaftlich inhaltliche Sicht, welche die Validität angebotener Informationsinhalte sowie den Grad der Unterstützung von Analyse- und Entscheidungsprozessen untersucht. Im Bereich *Technik* werden Komponenten und Architekturen betrachtet, indem insbesondere Aspekte wie die Flexibilität des Systementwurfs, die Qualität der IT-Lösung und das Ausmaß der Standardisierung analysiert werden. Der Bereich *Organisation* widmet sich der Einbettung des BI-Systems in die Aufbaustrukturen und Ablaufprozesse der Unternehmung. Es werden die Aspekte Wirtschaftlichkeit sowie Institutionalisierungs- und Formalisierungsgrad der BI-Prozesse betrachtet. Das biMM identifiziert insgesamt fünf Entwicklungsstufen, die durch Kriterien der drei

Schlüsselbereiche charakterisiert sind. Die erste Stufe ist ein *vordefiniertes Berichtswesen* (1). Der weitere Entwicklungspfad verläuft über *BI pro Fachbereich* (2), *unternehmensweite BI* (3) und *erweiterte Entscheidungsunterstützung* (4) bis hin zum *aktiven Wissensmanagement* (5).

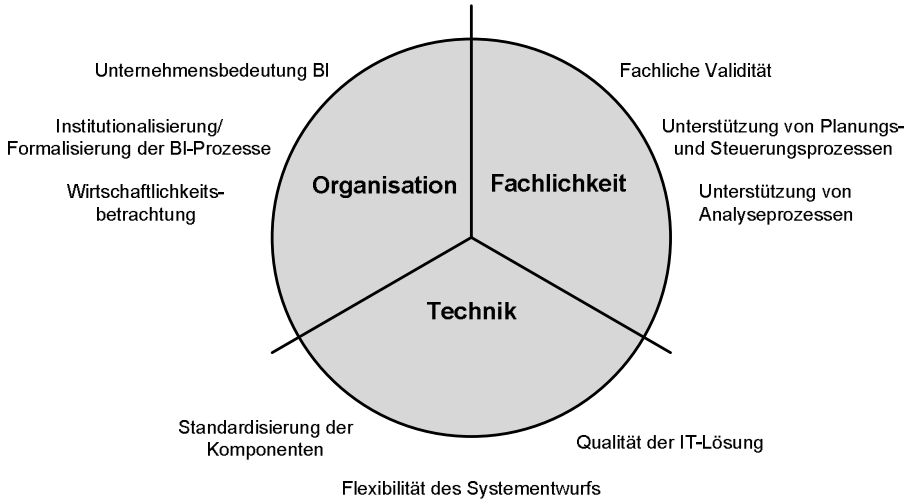


Abbildung 2: Untersuchungsfelder des BI-Reifegrads im biMM; in Anlehnung an [CG04]

Ein weiteres Modell, das Reifegrade von Business Intelligence-Systemen beschreibt, ist das *Business Intelligence Maturity Model* nach [Ec06]. Dieses Reifegradmodell identifiziert sechs Entwicklungsstufen, die durch verschiedene Merkmale charakterisiert sind, wie z. B. Anwendungsbereich, analytische Struktur, Wahrnehmung durch Führungskräfte, technische Plattform und Finanzierung. Die Bezeichnungen der Stufen orientieren sich am menschlichen Lebenszyklus und lauten *Prenatal* (1), *Infant* (2), *Child* (3), *Teenager* (4), *Adult* (5) und *Sage* (6). Die im Reifegradmodell abgebildete Glockenkurve signalisiert, dass die meisten Organisationen heute die Stufen 3 und 4 erreicht haben (vgl. Abbildung 3) [Ec06]. Auf dem sechsstufigen Entwicklungspfad identifiziert [Ec06] zwei besonders schwer zu überschreitende Hürden, die er *Gulf* und *Chasm* nennt. Um die erste Hürde zu überqueren muss ein Wechsel von individuellen Tabellenkalkulationen und Desktop-Datenbanken auf eine standardisierte Informationsbereitstellung erfolgen. Diese führt zu verbesserten Möglichkeiten der Nutzer, auf Informationen zuzugreifen, sie zu analysieren und entsprechend zu handeln. Um die zweite Hürde zu überqueren, muss die Unternehmensführung Business Intelligence als strategisch bedeutend und geschäftskritisch identifizieren. Aufgabe der IT-Abteilung ist dann, die durch interne Entwicklungen, Fusionen und Akquisitionen entstandenen einzelnen analytischen Systeme zu einem *Enterprise Data Warehouse* zu konsolidieren.

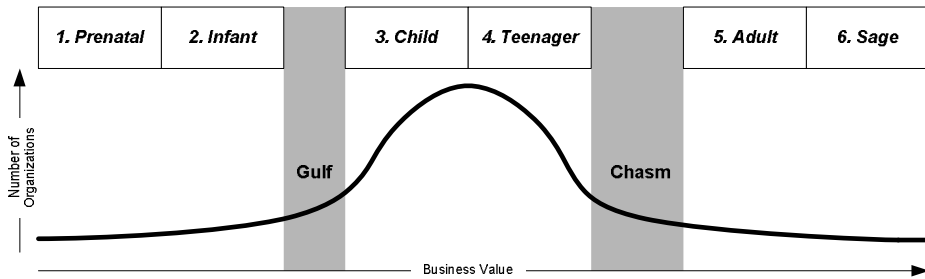


Abbildung 3: Glockenkurve des Business Intelligence Maturity Model; in Anlehnung an [Ec06]

[WK02] präsentieren ein Reifegradmodell zur Beurteilung von Performance Measurement-Systemen, die für ein Set von finanziellen und nicht-finanziellen Indikatoren Performance-relevante Daten aus verschiedensten Quellen sammeln, die Werte dieser Indikatoren mit historischen und angestrebten Werten vergleichen und die Ergebnisse an die beteiligten Personen verteilen. Sie identifizieren insgesamt vier Stufen beginnend bei *ad-hoc* über *heranwachsend* und *erwachsen* bis hin zu *reif*. Die Charakteristika der vier Stufen werden anhand der sechs Dimensionen *berücksichtigte Performance Measurement-Indikatoren*, *Datensammlung*, *Datenspeicherung*, *Kommunikation der Performance-Resultate*, *Nutzung der Performance-Resultate* und *Qualität des Performance Measurement-Prozesses* beschrieben [WK02].

Die vorgestellten Reifegradmodelle dienen als Grundlage für das in dieser Arbeit präsentierte Reifegradmodell, da sie jeweils Teilaspekte eines IT Performance Measurement, das durch BI-Systeme unterstützt wird, bereits abdeckten. Sie lieferten Ausgangslösungen zur Strukturierung des Reifegradmodells durch ihre Stufeneinteilung sowie durch die Differenzierung von Reifegraden mit Hilfe von Schlüsselbereichen und Kriterien.

3 Entwicklung des IT Performance Measurement Maturity Model

3.1 Überblick über das iterative Vorgehen

Die Entwicklung eines Reifegradmodells lässt sich als Gegenstand von *Design Science Research* nach [He04] einordnen. Design Science strebt eine Verbesserung der Problemlösungsfähigkeit durch die Schaffung innovativer IT-Artefakte an. Die im Rahmen von Design Science geschaffenen Artefakte umfassen sämtliche Elemente, die bei der Entwicklung von Informationssystemen zu berücksichtigen sind. Hierbei handelt es sich insbesondere um Sprachkonstrukte, Modelle, Methoden und Implementierungen. Das ITPM³ stellt ein solches Artefakt dar.

Nach den in [He04] formulierten Richtlinien für Design Science, soll Design Science in einer mehrfachen Wiederholung zweier aufeinander folgender Schritte erfolgen. Im

ersten Schritt wird das Artefakt entwickelt, im zweiten Schritt wird es auf seinen Problemlösungsbeitrag hin evaluiert. Dem Bild einer Spirale folgend soll so nach Möglichkeit ein immer höheres Niveau der Problemlösungsqualität und ihres empirischen Nachweises erreicht werden.

In Anlehnung an diese forschungsmethodischen Vorgaben gliederte sich der Prozess zur Entwicklung des Reifegradmodells in drei wesentliche Phasen (vgl. Abbildung 4). Jede Phase schloss mit der Explikation einer Version des Reifegradmodells ab. Zu Beginn der zweiten und dritten Phasen erfolgte jeweils eine Überprüfung der vorliegenden Version, die eine Überarbeitung des Modells zur Folge hatte. Der prinzipiell stetigen Fortsetzbarkeit dieses Prinzips entsprechend stehen weitere Validierungen der hier präsentierten Version des Reifegradmodells noch aus. Im Folgenden werden die einzelnen Phasen detailliert dargestellt.

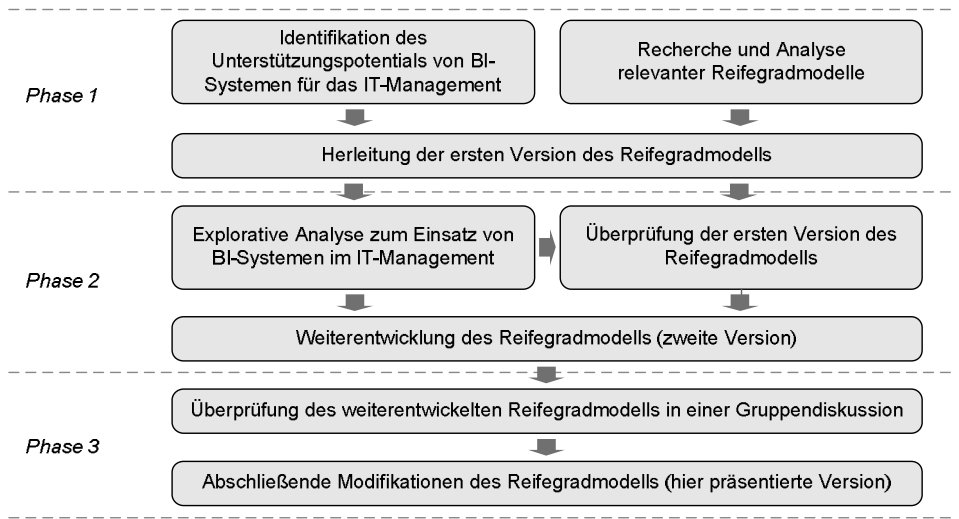


Abbildung 4: Entwicklungsprozess des Reifegradmodells

3.2 Ergebnisse der ersten Entwicklungsphase

Die erste Phase befasste sich mit der Abgrenzung des Anwendungsbereichs des Reifegradmodells und einer Analyse relevanter Reifegradmodelle aus den Bereichen Business Intelligence und IT-Management (vgl. Abschnitt 2). Es wurden Anforderungen an das Modell aus formaler und inhaltlicher Sicht aufgestellt (vgl. Tabelle 1).

Anforderungen aus inhaltlicher Sicht	Anforderungen aus formaler Sicht
<ul style="list-style-type: none">• Identifikation von sinnvollen Entwicklungsstufen bei der Ausgestaltung des IT Performance Measurement unter besonderer Berücksichtigung des Einsatzes von Business Intelligence-Systemen.• Eignung als Hilfestellung bei der Analyse der Ist-Situation sowie bei der Herleitung von zukünftigen Zielzuständen und Verbesserungsmaßnahmen unter Berücksichtigung der Ist-Situation.• Gültigkeit des Modells möglichst für sämtliche IT-Organisationen unabhängig von Größe und Branche des betrachteten Unternehmens.• Berücksichtigung aller diversen relevanten Aspekte auch abseits der Technik.• Integration inhaltlich passender Elemente aus anderen bewährten Modellen.• Praxisnähe.	<ul style="list-style-type: none">• Verwendung natürlicher Sprache, um ein grundsätzliches Verständnis des Modells ohne zusätzlichen Aufwand durch Erlernen von abweichenden Methoden oder Techniken zu gewährleisten.• Orientierung an Strukturen bekannter Reifegradmodelle, da dies das Verständnis und die Anwendung des Modells zusätzlich erleichtern kann.• Sinnvolle Anzahl an Entwicklungsstufen.• Ausreichender aber nicht zu geringer Umfang der Beschreibungen von Entwicklungsstufen.• Vorgehensbeschreibung zur Bestimmung von Reifegraden, um durch eine einheitliche Anwendung von Richtlinien und Standards eine Vergleichbarkeit der Ergebnisse zu gewährleisten.

Tabelle 1: Anforderungen an das Reifegradmodell

Im Anschluss folgte die Konstruktion einer ersten Version des Reifegradmodells, die auf bestehenden Reifegradmodellen basierte und vier Entwicklungsstufen umfasste. Als erste Stufe wurde das *fragmentierte IT-Berichtswesen* identifiziert. Die zweite Stufe bildete das *konsolidierte IT-Berichtswesen mit Hilfe von Tabellenkalkulation*, gefolgt von der dritten Stufe, dem *IT Performance Dashboard*. Die vierte und höchste Stufe war das *Informationsportal für das IT-Management*. Das Reifegradmodell gliederte sich in Anlehnung an das biMM von [CG04] in die Dimensionen *Fachlichkeit*, *Organisation* und *Technik*. Die Dimension Fachlichkeit befasste sich mit der inhaltlichen Relevanz der umgesetzten IT Performance Measurement-Lösung für das IT-Management. Die Dimension Organisation untersuchte die Einbettung der Lösung in die Aufbaustrukturen und Abläufe der IT-Organisation sowie ihre Integration in unternehmensweite Konzepte. Die Dimension Technik betrachtete die verwendeten Komponenten und Architekturen. Diese Dimensionen wurden durch jeweils fünf Kriterien näher charakterisiert.

3.3 Ergebnisse der zweiten Entwicklungsphase

Die zweite Phase umfasste eine explorative Analyse zum Einsatz von BI-Systemen zur Unterstützung des IT-Managements sowie eine kritische Überprüfung der ersten Version des Reifegradmodells. Dies geschah im ersten Halbjahr 2007 im Rahmen von semi-strukturierten Interviews mit Vertretern aus dem IT-Management deutscher Unternehmen (Gesprächsdauer jeweils ca. 45 bis 60 Minuten; vgl. Interviewleitfaden im Anhang). Mit Hilfe dieser Interviews wurde exemplarisch erhoben, inwieweit und in welcher Form BI-Systeme aktuell zur Unterstützung des IT-Managements eingesetzt werden. Außerdem wurde die erste Version des ITPM³ kritisch geprüft sowie Sichtweisen und Anforderungen aus der Praxis erfragt, um diese in die Weiterentwicklung des Reifegradmodells einfließen zu lassen.

GP	Position des GP	Branche/ Beschreibung des Unternehmens	Umsatz (in € pro Jahr)	Mitarbeiter	IT-Ausgaben (in € pro Jahr)	IT-Mitarbeiter
A	Leiter Rechnungswesen	IT-Dienstleister einer Finanzdienstleister- gruppe	13000 Mio. (Gesamtbeiträge der Gruppe)	19.000 (Mitarbeiter der Gruppe)	400 Mio. (Umsatz des IT- Dienstleisters)	735 (Mitarbeiter des IT-Dienstl.)
B	Leiter IT-Projekt- management	Kabelnetzbetreiber	1.000 Mio.	2.700	-	140
C	CIO	IT-Dienstleister eines Industriekonzerns	5.000 Mio.	43.000	-	-
D	Director International Telecommunications	IT-Dienstleister im Gesundheitswesen	10.000 Mio. (Mutterkon- zern)	100.000 (Mutterkon- zern)	-	200 (Mitarbeiter des IT-Dienstl.)
E	Mitarbeiter im Bereich IT-Strategie und Organisations- entwicklung	Hausgerätehersteller	8.500 Mio.	35.000	170 Mio.	850
F	Head of Project Management	Touristik	7.800 Mio.	20.000	65 Mio.	200
G	CIO	Mittelständisches Unternehmen in der Elektroindustrie	430 Mio.	3.000	10 Mio.	65
H	Head of IT Logistics Performance & Quality Management	Logistik	23.000 Mio.	160.000	-	2.000
I, J	Leiter/ Mitarbeiter IT- Controlling	Zulieferer der Automobilindustrie	15.000 Mio.	85.000	200 Mio.	900

Tabelle 2: Stichprobe der explorativen Analyse¹

Die Stichprobe der befragten Personen umfasste insgesamt 10 Gesprächspartner (GP) aus 9 Unternehmen (vgl. Tabelle 2; anonymisiert), die alle über ein abgeschlossenes Studium in den Bereichen der Wirtschafts- oder Naturwissenschaften verfügten (bspw. BWL und VWL, Wirtschaftsingenieurwesen, Nachrichtentechnik oder Informatik). Die Stichprobe war in der Hinsicht homogen, dass alle Gesprächspartner zum Zeitpunkt der Interviews im IT-Management oder IT-Controlling beruflich tätig waren und diesbezüglich langjährige Erfahrung (bis zu 25 Jahre) aufwiesen. Innerhalb dieses Bereichs unterschieden sich jedoch ihre Aufgaben und Verantwortungsbereiche. Einzelne waren in der Position des Chief Information Officer (CIO), während andere das IT-Management z. B. aus der Sicht des Rechnungswesens betrachteten. Des Weiteren repräsentierten die befragten Personen ein weites Spektrum unterschiedlicher Unternehmen in Hinblick auf Größe und Branche, angefangen bei mittelständischen Unternehmen über IT-Dienstleister großer Industriekonglomerate bis hin zu Touristik-Unternehmen. Hierdurch ergab sich zugleich eine gewisse, durchaus gewünschte, Heterogenität der Stichprobe, die es ermöglichen sollte, möglichst allgemeingültige Schlussfolgerungen abzuleiten.

¹ Die Angaben zu Umsatz, Mitarbeiterzahlen u. ä. stammen aus den einzelnen Gesprächen oder von Internet-auftritten der jeweiligen Unternehmen bzw. Drittanbietern. Die angegebenen Werte sollen lediglich Größenordnungen der betrachteten Unternehmen vermitteln. Die Werte sind jeweils gerundet und können sich aufgrund der heterogenen Datengrundlage auf unterschiedliche Jahre beziehen. Die Angaben zu Gesprächspartnern und Unternehmen wurden auf Wunsch der Interviewten anonymisiert.

Die Ergebnisse der Interviews zeigen, dass ein IT Performance Measurement im Sinne eines umfassenden Blicks auf die Effizienz und Effektivität der IT selten durch die eingesetzten computerbasierten Werkzeuge ermöglicht wird. Benötigte Informationen werden – wenn sie überhaupt verfügbar sind – in verschiedenen Systemen gehalten, durch unterschiedliche Berichte verbreitet und stehen isoliert nebeneinander. Die finanziellen Daten werden zumeist in einem unternehmensweiten ERP- oder Finanzinformationssystem (wie z. B. einem SAP-System) vorgehalten. Basierend auf diesen Daten lassen sich Berichte mit entsprechenden Finanzkennzahlen automatisiert generieren. Die nicht-finanziellen Kennzahlen hingegen sind in verschiedenen anderen operativen Systemen gespeichert und werden häufig mit Hilfe von manuell erstellten Tabellenkalkulationen o. ä. berichtet. Es existiert in den meisten Unternehmen keine integrierte Datenbasis für den IT-Bereich, auf die ein einheitliches Berichtswesen aufbauen könnte und eine automatisierte Verknüpfung von finanziellen mit nicht-finanziellen Kennzahlen möglich wäre. Ebenso fehlen häufig einheitliche Kennzahlensysteme oder Konzepte wie eine IT Balanced Scorecard, die finanzielle und nicht-finanzielle Aspekte verbinden und zusätzlich eine Orientierung an Unternehmenszielen gewährleisten.

Gleichwohl lässt sich erkennen, dass in fast allen untersuchten Unternehmen Initiativen bereits gestartet wurden oder in Planung sind, um deren IT Performance Measurement zu verbessern (vgl. Tabelle 3). Insbesondere wurde die Entwicklung einheitlicher Kennzahlensysteme genannt, die die finanziellen und nicht-finanziellen Aspekte der IT integrieren und somit eine verbesserte Transparenz über die Effizienz und Effektivität gewährleisten sollen. Darüber hinaus streben vor allem die größeren Unternehmen eine zunehmende Automatisierung mit Hilfe von Business Intelligence-Werkzeugen an, bspw. mit Data Warehouses und entsprechend darauf aufbauenden Anwendungen. Hierdurch soll manueller Aufwand und das Arbeiten mit Tabellenkalkulationen vermindert werden.

Basierend auf den hinzugewonnenen Erkenntnissen wurde das Reifegradmodell grundlegend überarbeitet. Das IT Performance Measurement wurde in Anlehnung an das prozessorientierte Verständnis von Business Intelligence nach [GG00] als ein *analytischer Prozess* definiert, der fragmentierte interne und externe Daten in Transparenz über die Effizienz und Effektivität der eigenen Informationsinfrastruktur sowie in handlungsgerichtetes Wissen über ihre Positionierung, Potenziale und Perspektiven transformiert. Hierbei umfasst der Begriff *Informationsinfrastruktur* sämtliche Einrichtungen, Mittel und Maßnahmen zur Produktion, Verbreitung und Nutzung von Information in Organisationen [HL05]. Darüber hinaus wurde die Stufeneinteilung modifiziert und umfasste nun analog zu den CobiT Maturity Models sechs Reifegrade von *nicht existent* (0) bis *optimiert* (5). Die Stufe 0 stand für ein gänzlich Fehlen eines IT Performance Measurement-Ansatzes. Im Zuge der Weiterentwicklung wurden die Kriterien der drei Dimensionen ebenfalls überprüft und modifiziert.

G P	Bisherige Ansätze	Aktuelle und geplante Initiativen
A	<ul style="list-style-type: none">• Informationen über IT-Kosten werden aus einem SAP-System bezogen.• Kosten- und Leistungsrechnung wurde bereits auf eine Profit Center-orientierte Kostenstellensteuerung umgestellt.• Produkte werden als komplette Systeme nach geschäftsgetriebenen Mengentreibern abgerechnet.	<ul style="list-style-type: none">• Entwicklung eines übergeordneten, einheitlichen und unternehmensweiten Kennzahlensystems, das nicht-finanzielle Faktoren integriert.• Aufbau eines integrierten Datenspeichers der neben finanziellen Daten auch aggregierte technische Daten und Mengentreiber bereitstellen soll.• Integration von Simulationstools, um alternative Szenarien zu rechnen.
B	<ul style="list-style-type: none">• Finanzdaten werden aus einem SAP-System bezogen.• Weitere Informationen kommen aus diversen operativen Systemen wie z. B. einem Project Management Tool.• Aufbereitung und Bereitstellung von Informationen geschieht zumeist mit Excel.	<ul style="list-style-type: none">• Entwicklung eines integrierten Kennzahlensystems, das bisher erhobene Kennzahlen verbindet.
C	<ul style="list-style-type: none">• Es befinden sich zurzeit zwei Data Warehouses im Einsatz. Ersteres umfasst bisher ausschließlich Finanzinformationen. Das Zweite speichert operative Daten.• Einsatz von Tools des Anbieters Mercury sowie von Eigenentwicklungen zur Unterstützung des IT-Managements.	<ul style="list-style-type: none">• Entwicklung eines einheitlichen Kennzahlensystems• Zunehmende Prozessstandardisierung.• Erweiterung des ersten Data Warehouse um aggregierte Kennzahlen aus dem operativen Bereich.• Zunehmende Verdrängung von Excel als Werkzeug für Auswertungen und Planungen.
D	<ul style="list-style-type: none">• Bisher existiert kein einheitlicher Ansatz für das IT Performance Measurement.• Finanzielle Daten werden durch ein SAP-System sowie einer darauf aufsetzenden Anwendung zur Berichtsgenerierung bereitgestellt.• Auf Vorstandsebene wird häufig Excel genutzt, um Daten aus verschiedenen Systemen zu verknüpfen.	<ul style="list-style-type: none">• Ein Ticket-System befindet sich in der Einführung, das den Einsatz von Excel in diesem Bereich ablösen soll.
E	<ul style="list-style-type: none">• Einsatz eines webbasierten, teilautomatisierten Performance Monitors, der IT-Kennzahlen in vier Dimensionen strukturiert. Diese sind als Dokumente statisch verlinkt.	<ul style="list-style-type: none">• Integration von CobiT-Performance-Indikatoren.• Durchführung von Assessments mit Hilfe von CobiT geplant, um Reife von Prozessen zu bestimmen.
F	<ul style="list-style-type: none">• Einsatz eines Webportals um einzelne Projekte zu managen.• Finanzdaten liegen in einem SAP-System.	<ul style="list-style-type: none">• Stärkere Automatisierung des Portfoliomanagements und der Auslastungsplanung, die zurzeit jeweils noch auf Excel-Listen basieren.
G	<ul style="list-style-type: none">• Einsatz eines an ITIL orientierten Kennzahlensystems (IT Service Report).• Umsetzung des IT Service Reports mit Excel• Teilweise automatisierte Dateneinflege mit Hilfe von Makros.• Assessment interner Prozesse nach dem Capability Maturity Model.	<ul style="list-style-type: none">• Erweiterung des IT Service Reports um weitere Performance-Indikatoren.• Vollständig automatisierte Anbindung der Datenquellen.
H	<ul style="list-style-type: none">• CIO Dashboard, das auf einem definierten Prozessmodell aufbaut und sowohl strategische als auch operative Prozesse umfasst.	<ul style="list-style-type: none">• Frage der Tool-Unterstützung des CIO Dashboard soll neu diskutiert werden.
I, J	<ul style="list-style-type: none">• Standardisiertes Berichtswesen der IT-Kosten aufgeschlüsselt nach 20 Standard-Prozessen basierend auf einem SAP-System.• Bereitstellung operativer Informationen aus dezentralen Datenbeständen mittels Excel-Dateien und manuell erstellten Berichten.	<ul style="list-style-type: none">• Verknüpfung von finanziellen und nicht-finanziellen Kennzahlen in einem Kennzahlensystem.• Umsetzung des Konzepts mit einem Data Warehouse, in das benötigte Daten automatisiert aus den Quellsystemen geladen werden. Vermeidung manueller Zwischenlösungen.

Tabelle 3: Übersicht über bisherige Ansätze sowie aktuelle und geplante Initiativen

3.4 Ergebnisse der dritten Entwicklungsphase

In der dritten Phase wurde das weiterentwickelte Reifegradmodell erneut einer konsensbasierten Validierung unterzogen, indem es in einer Gruppendiskussion mit drei IT-Beratern der *Deloitte Consulting GmbH* kritisch geprüft wurde. Wiederum wurde das Reifegradmodell an hinzugewonnene Erkenntnisse angepasst. Durch abschließende Modifikationen, die insbesondere eine inhaltliche Schärfung der Dimension Fachlichkeit und ihrer Kriterien bezweckten, entstand die vorläufig letzte Version des ITPM³ (vgl. Abbildung 5).

0	Nicht existent	IT Performance-Intransparenz
1	Initial	Ad-hoc IT Performance Measurement
2	Wiederholbar	Fragmentiertes, operativ ausgerichtetes IT Performance Measurement
3	Definiert	Definierter IT Performance Measurement-Prozess, der auf ein mit Desktop-Werkzeugen unterstütztes, konsolidiertes IT-Berichtswesen zurückgreift.
4	Gemanagt	Formalisierter und automatisierter IT Performance Measurement-Prozess, der durch ein leistungsfähiges IT Performance Dashboard unterstützt wird.
5	Optimiert	Vernetzter IT Performance Measurement-Prozess, der durch unternehmensweit integrierte Business-Intelligence-Werkzeuge unterstützt wird.

Abbildung 5: Entwicklungsstufen des ITPM³

Die Reifegrade des Modells beschreiben den analytischen BI-Prozess, der fragmentierte interne und externe Daten in handlungsgerichtete Informationen über die Effizienz und Effektivität der eigenen Informationsinfrastruktur transformiert, von (0) *nicht existent* bis (5) *optimiert*. Der skizzierte Entwicklungspfad beginnt bei einer vollkommenen IT Performance-Intransparenz und endet in der höchsten Stufe bei einem vernetzten IT Performance Measurement-Prozess, der durch unternehmensweit integrierte BI-Werkzeuge unterstützt wird. Die sechs Stufen lassen sich wie folgt beschreiben:

0. *Nicht existent*: Das IT-Management führt kein eigenständiges Monitoring von IT-Projekten, der Informationsinfrastruktur oder von IT-Prozessen durch. Nützliche, aktuelle und akkurate Berichte über die IT Performance sind deshalb nicht verfügbar. Es herrscht eine *IT Performance-Intransparenz*. Die Stufe 0 steht für ein gänzlich fehlendes IT Performance Measurement-Ansatzes. Daher wird diese Stufe im Reifegradmodell nicht näher auf Ebene der Kriterien und Dimensionen beschrieben.
1. *Initial*: Auf dieser Stufe ist der Bedarf des IT Performance Measurement zwar erkannt, jedoch existiert kein entsprechendes Konzept. Es fehlen standardisierte Prozesse für die Sammlung und Auswertung von Daten. Dies geschieht stattdessen durch manuelle, wenig organisierte Ad-hoc-Prozesse. Die Bereitstellung und Auswertung von Kennzahlen erfolgt von Fall zu Fall für das jeweilige betrachtete Objekt und zumeist als Reaktion auf Zwischenfälle. Ziele sind nicht klar und deren Erfüllung wird nicht gemessen. Das Rechnungswesen überwacht grundlegende Finanzkennzahlen der IT. Dieser Reifegrad lässt sich als *Ad-hoc IT Performance Measurement* zusammenfassen.
2. *Wiederholbar*: Der Fokus des IT Performance Measurement liegt auf Effizienz von Entwicklung und Betrieb von IT-Systemen. Es existieren diesbezüglich Methoden und Techniken zur Sammlung und Auswertung von Daten, diese werden jedoch un-

einheitlich angewandt. Datenbank-gestützte Monitoring-Tools sowie Standard-Berichte für einzelne Betrachtungsobjekte werden genutzt. Dennoch ergibt sich bei der Weiterverarbeitung von Daten aus den verschiedenen Quellen in Tabellenkalkulationen ein hoher manueller Aufwand. Die Interpretation der Ergebnisse hängt von einzelnen Schlüsselpersonen ab. Zusammenfassend handelt es sich hierbei um ein *fragmentiertes, operativ ausgerichtetes IT Performance Measurement*.

3. *Definiert*: Ein erstes Konzept für das IT Performance Measurement wird umgesetzt, wie z. B. eine IT Balanced Scorecard. Entsprechende finanzielle und nicht-finanzielle Messgrößen sind definiert. Das IT-Management führt institutionalisierte Standard-Monitoring-Prozesse innerhalb der IT-Organisation ein. In diesem Zusammenhang werden automatisierte Tools für die Sammlung und Überwachung operativer Informationen zu Anwendungen, Systemen und Prozessen genutzt. Die möglicherweise notwendige manuelle Dateneinpfege ist formalisiert. Auf dieser Entwicklungsstufe existiert somit ein *definierter IT Performance Measurement-Prozess, der auf ein mit Desktop-Werkzeugen unterstütztes, konsolidiertes IT-Berichtswesen zurückgreift*.
4. *Gemanagt*: Ein IT Performance Measurement-Konzept wurde in Abstimmung mit der Unternehmensführung entwickelt, so dass sich die Messgrößen der IT-Funktion an den Unternehmenszielen orientieren. Basierend auf einer entsprechenden Prozessdokumentation hat das IT-Management Toleranzgrenzen für die Performance von IT-Prozessen definiert. Alle Metriken sämtlicher IT-Projekte und IT-Prozesse sind Teil eines integrierten Kennzahlensystems, dem einheitliche Definitionen zugrunde liegen. Die Sammlung und Überwachung operativer Daten zu Anwendungen, Systemen und Prozessen ist automatisiert. Die Daten werden in einen integrierten Datenspeicher geladen, der Basis für leistungsstarke Analysewerkzeuge ist. Mit dem Erreichen dieses Reifegrads existiert ein *formalisierter und automatisierter IT Performance Measurement-Prozess, der durch ein leistungsfähiges IT Performance Dashboard unterstützt wird*.
5. *Optimiert*: Das IT Performance Measurement-Konzept ist funktions- und bereichsübergreifend in das Performance Measurement-System des gesamten Unternehmens integriert. Geschäftsgetriebene Kennzahlen werden zur Performance-Messung benutzt. Alle Monitoring-Prozesse sind optimiert. Abweichungen werden umfassend und durchgängig ausgewertet sowie deren Ursachen analysiert. Benchmarking ist formalisiert und in den Performance Measurement-Prozess integriert. Die Sammlung und Überwachung operativer Daten zu Anwendungen, Systemen und Prozessen ist automatisiert. Die Daten werden in einen integrierten Datenspeicher geladen, der Basis für leistungsstarke, unternehmensweit standardisierte Analysewerkzeuge ist. Zusammenfassend beschreibt dieser Reifegrad einen *vernetzten IT Performance Measurement-Prozess, der durch unternehmensweit integrierte Business Intelligence-Werkzeuge unterstützt wird*.

Um eine differenzierte Analyse des IT Performance Measurement zu ermöglichen, umfasst das Reifegradmodell drei Dimensionen, die jeweils durch fünf Kriterien näher charakterisiert werden (vgl. Abbildung 6). Die Bezeichnung der Reifegrade sowie die Prozessorientierung lehnen sich an die CobiT Maturity Models an. Die Strukturierung in die Dimensionen *Fachlichkeit*, *Organisation* und *Technik* stammt aus dem Business Intelligence Maturity Model von [CG04]. Zusammen mit dem Business Intelligence Maturity Model nach [Ec06] diente dies außerdem insbesondere in der Dimension *Technik* zur Charakterisierung der Stufen. Einzelne Kriterien wie *Wirtschaftlichkeitsbetrachtung* wurden ebenfalls übernommen und angepasst. Weitere Kriterien lehnen sich an die Attribute der Reife des generischen *CobiT* Reifegradmodells an (wie z. B. *Richtlinien*, *Standards und Verfahren*). Bedeutenden Einfluss auf die grundsätzlichen Charakterisierungen der Stufen hatten außerdem das Reifegradmodell für den CobiT-Prozess *ME1* [ITG05] sowie das *IT-BSC Maturity Model* von [GS01].

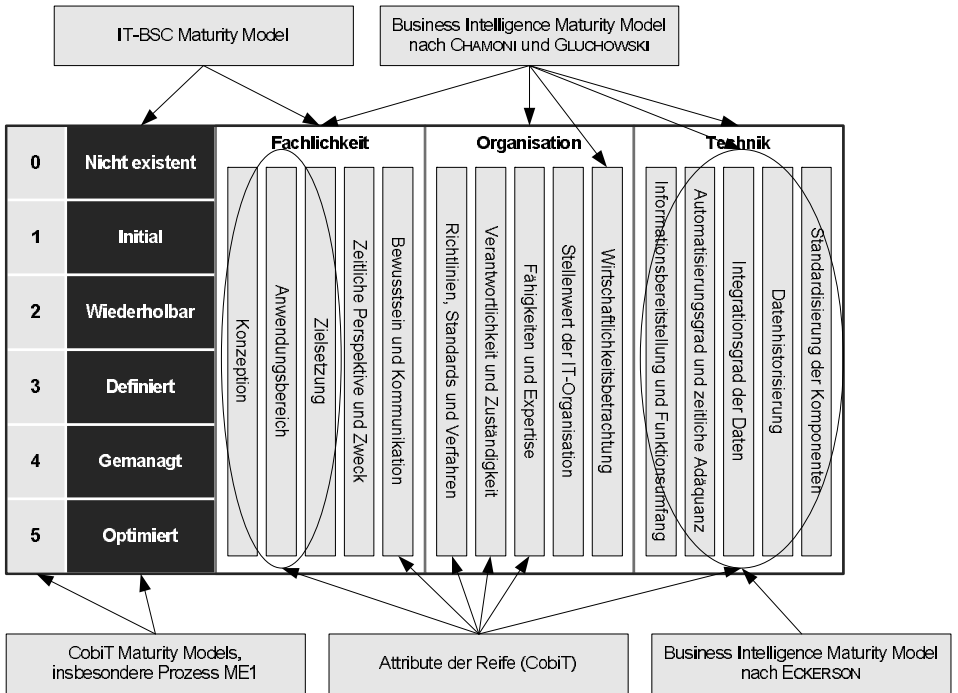


Abbildung 6: Übersicht über die Einflüsse anderer Reifegradmodelle

Die Dimensionen *Fachlichkeit*, *Organisation* und *Technik* werden durch jeweils fünf Kriterien näher charakterisiert, die einzelne spezifische Fragestellungen ihrer Dimension betrachten (vgl. Tabelle 4). Die Dimension *Fachlichkeit* befasst sich mit der inhaltlichen Relevanz der umgesetzten IT Performance Measurement-Lösung für das IT-Management. Die Dimension *Organisation* untersucht die Einbettung der Lösung in die Aufbaustrukturen und Abläufe der IT-Organisation sowie deren Integration in unternehmensweite Konzepte. In der Dimension *Technik* werden die verwendeten Komponenten und Architekturen betrachtet.

Dimension	Kriterium	Fragestellung
Fachlichkeit	Konzeption	Inwieweit existiert ein Konzept für das IT Performance Measurement?
	Anwendungsbereich	Inwieweit werden relevante Bereiche durch das IT-Kennzahlensystem abgedeckt?
	Zielsetzung	Inwieweit sind Kennzahlen unternehmensweit vereinheitlicht und inwieweit orientiert sich deren Zielvorgabe an Unternehmenszielen sowie externen Benchmarking-Referenzen?
	Zeitliche Perspektive und Zweck	Inwieweit wird versucht, durch das IT Performance Measurement die zukünftige Leistungsfähigkeit der IT zu verbessern?
	Bewusstsein und Kommunikation	Inwieweit ist der Bedarf für ein IT Performance Measurement erkannt und wie werden Performance-Resultate kommuniziert?
Organisation	Richtlinien, Standards und Verfahren	Inwieweit sind die Abläufe des IT Performance Measurement durch Richtlinien, Standards und Verfahren institutionalisiert und formalisiert?
	Verantwortlichkeiten und Zuständigkeit	Inwieweit sind Zuständigkeiten und Verantwortlichkeiten festgelegt und anerkannt?
	Fähigkeiten und Expertise	Inwieweit sind entsprechende Fähigkeiten und Expertise im Unternehmen vorhanden?
	Stellenwert der IT-Organisation	Inwieweit ist die IT-Organisation und somit auch das IT Performance Measurement in der Organisation von Bedeutung?
	Wirtschaftlichkeitsbetrachtung	Inwieweit werden Kosten und Nutzen der BI-Werkzeuge zur Unterstützung des IT Performance Measurement analysiert?
Technik	Informationsbereitstellung und Funktionsumfang	Inwieweit entsprechen die bereitgestellten Informationen und Analysefunktionen den Bedürfnissen des Nutzers?
	Automatisierung und zeitliche Adäquanz	Inwieweit sind Datensammlung, -speicherung und -aufbereitung sowie Informationsbereitstellung automatisiert und welche Aktualität wird dadurch gewährleistet?
	Integrationsgrad der Daten	Inwieweit ist die Datenbasis konsistent und integriert?
	Datenhistorisierung	Inwieweit ist die Archivierung automatisiert und wie gut sind historische Datenstände verfügbar?
	Standardisierung der Komponenten	Inwieweit wird innerhalb der IT-Organisation und des Gesamtunternehmens ein einheitliches Set von Werkzeugen genutzt?

Tabelle 4: Kriterien und ihre Fragestellungen

Stufe	Fachlichkeit	Organisation	Technik
0	-	-	-
1	<ul style="list-style-type: none"> Es existiert kein Konzept für das IT Performance Measurement. Ziele sind nicht klar. Folglich kann das Erreichen dieser Ziele nicht gemessen werden. Beschränkung auf grundlegende Finanzkennzahlen. 	<ul style="list-style-type: none"> Prozesse und Richtlinien sind nicht definiert. Zuständigkeiten und Verantwortlichkeiten sind nicht festgelegt. Die IT-Organisation wird als Cost Center angesehen. 	<ul style="list-style-type: none"> Einige computergestützte Werkzeuge existieren, hierbei handelt es sich vor allem um Desktop-Tools. Es existiert keine geplante Herangehensweise für den Tool-Einsatz.
2	<ul style="list-style-type: none"> Monitoring ist inkonsistent und auf Einzelbereiche konzentriert (wie z. B. einzelne Applikationen, IT-Prozesse oder Projekte). Einige Ziele werden gesetzt. Insbesondere finanzbezogene Kennzahlen sowie Kennzahlen zur Effizienz von Entwicklung und Betrieb sind entwickelt. 	<ul style="list-style-type: none"> Ähnliche allgemeine Prozesse entwickeln sich, aber basieren auf Intuition und individueller Expertise. Einzelpersonen vermuten deren Zuständigkeit. Die IT-Organisation wird als Cost Center angesehen. 	<ul style="list-style-type: none"> Allgemeine Ansätze für den Tool-Einsatz sind erkennbar, aber Lösungen werden durch Einzelpersonen entwickelt (z. B. in Form von wiederverwendbaren Tabellenkalkulations-Templates). Vorhandene Werkzeuge zum Monitoring der Informationsinfrastruktur werden unter Umständen nicht ausreichend genutzt.
3	<ul style="list-style-type: none"> Ideen der IT Balanced Scorecard werden aufgenommen und prototypisch umgesetzt. IT-spezifische, nicht-finanzielle und strategische Messgrößen, Messgrößen zur Kundenzufriedenheit und Service Levels sind definiert. Bei der Vorgabe von Zielwerten besteht eine Orientierung an den Unternehmenszielen. Ursachen für Abweichungen werden analysiert 	<ul style="list-style-type: none"> Der IT Performance Measurement-Prozess sowie Richtlinien und Verfahren für wesentliche Aktivitäten sind definiert und dokumentiert. Datenquellen und Verantwortlichkeiten für Dateneingabe sind festgelegt. Zuständigkeiten und Verantwortlichkeiten sind definiert, werden jedoch nicht vollständig befolgt. Die IT-Organisation agiert als Service Center. 	<ul style="list-style-type: none"> Werkzeuge werden zum Monitoring wichtiger Betrachtungsobjekte angewandt, sie sind jedoch nicht integriert. Der Einsatz von Werkzeugen hat das Ziel, zunehmend Prozesse zu formalisieren. Daten werden durch zumeist manuelle Eingabe in einer Übersicht zusammengetragen, die mit Hilfe von Desktop-Tools (wie z. B. Tabellenkalkulation) erstellt wurde.
4	<ul style="list-style-type: none"> Effizienz und Effektivität der IT werden gemessen und kommuniziert. Sie werden mit Unternehmenszielen und strategischen Plänen der IT-Organisation in Verbindung gesetzt. Es existiert ein gefestigtes, bewährtes und dokumentiertes Konzept. Sämtliche Kennzahlen sind Teil eines umfassenden, integrierten und wohl definierten Kennzahlensystems, das auf einheitliche Begrifflichkeiten aufbaut. 	<ul style="list-style-type: none"> Sämtliche Aspekte des Prozesses sind dokumentiert und wiederholbar. Zuständigkeiten und Verantwortlichkeiten sind akzeptiert. Die IT-Organisation agiert als Profit Center mit einem eigenständigen Produkt- und Dienstleistungsangebot. Adäquate und umfassende Wirtschaftlichkeitsbetrachtung ist notwendig und wird durchgeführt. 	<ul style="list-style-type: none"> Der Einsatz von Werkzeugen ist geplant und integriert. Operative Daten bzgl. der Performance von IT-Systemen und IT-Prozessen werden automatisiert in einen integrierten Datenpool geladen, der Basis für leistungsstarke Analysewerkzeuge ist. Die Werkzeuge werden eingesetzt, um den IT Performance Measurement-Prozess zu automatisieren.
5	<ul style="list-style-type: none"> Es besteht ein integriertes System zur Performancemessung, welches IT Performance mit Unternehmenszielen verbindet. Abdeckung sämtlicher relevanten Perspektiven des IT Performance Measurement Kennzahlen wurden vollständig und wohl überlegt aus der strategischen Ausrichtung des Unternehmens abgeleitet. 	<ul style="list-style-type: none"> Abläufe sind vollständig formalisiert und institutionalisiert. Die Prozessdokumentation wurde zu automatisierten Workflows weiterentwickelt. Zuständigkeiten und Verantwortlichkeiten sind akzeptiert. Es wird eine kontinuierliche Verbesserung des IT Performance Measurement-Prozesses angestrebt. IT gilt als Enabler für neue Geschäftsmodelle und Strategien. Adäquate und umfassende Wirtschaftlichkeitsbetrachtung ist notwendig und wird durchgeführt. 	<ul style="list-style-type: none"> Rückgriff auf standardisierte Werkzeug-Sets, die im gesamten Unternehmen genutzt werden. In die Datenbasis werden auch externe Datenquellen integriert (z. B. für Benchmarking). Werkzeuge sind vollständig mit anderen Werkzeugen integriert und ermöglichen eine über die IT-Organisation hinausgehende Analyse. Abweichungen von Zielvorgaben werden automatisch erkannt.

Tabelle 5: Beschreibung der Entwicklungsstufen nach Dimensionen

Charakteristische Ausprägungen der drei Dimensionen in den jeweiligen Entwicklungsstufen werden in Tabelle 5 zusammengefasst.

3.5 Anwendung des Reifegradmodells

Die Anwendung des Reifegradmodells zur Analyse der IT Performance Measurement-Lösungen ist in zwei Varianten möglich. Die erste Variante geht *bottom-up* vor und beginnt mit einer umfassenden Beurteilung anhand sämtlicher Kriterien. Jedem einzelnen Kriterium wird hierbei der maximal mögliche Reifegrad zugewiesen, der dem IT Performance Measurement-Ansatz des betrachteten Unternehmens entspricht. Auf dieser Ebene ergibt sich dadurch ein differenziertes Reife-Profil. Anschließend wird über die Ebene der Dimensionen bis hin zur obersten Ebene aggregiert.

Die alternative Variante geht in der entgegengesetzten Reihenfolge vor (*top-down*). Hierbei beschränkt sich die Analyse des Ist-Zustands zunächst auf einen Abgleich mit den sechs Stufen des Reifegradmodells, ohne die Dimensionen oder Kriterien zu betrachten. Diese werden nur bei Bedarf schrittweise hinzugezogen, um die Einordnung auf eine der Stufen abzusichern.

Die an die Ermittlung des Reifegrades sich anschließende Ableitung von Verbesserungsmaßnahmen bedarf der vorherigen individuellen Festlegung eines Zielzustands. Dieser ist mit dem Ist-Zustand zu vergleichen, um mögliche Abweichungen zu bestimmen. Mit Hilfe der Beschreibungen des Reifegradmodells je Kriterium und Entwicklungsstufe lassen sich Handlungsempfehlungen ableiten, um diese Abweichungen zu eliminieren. Hierdurch kann ein wesentlicher Beitrag zur konzeptionellen sowie technischen Fortentwicklung des IT Performance Measurement in Unternehmen und anderen Organisation geleistet werden.

4 Fazit und Ausblick

Im Rahmen der Entwicklung des Reifegradmodells wurde eine explorative Analyse durchgeführt, die zeigte, dass viele Unternehmen im Bereich des IT Performance Measurement Handlungsbedarf verspüren und Initiativen zur Verbesserung gestartet haben. Einerseits stehen hierbei konzeptionelle Fortschritte wie die Integration von bisher fragmentierten Ansätzen durch die Entwicklung eines einheitlichen Kennzahlensystems für die IT-Organisation im Vordergrund. Andererseits wird insbesondere die Automatisierung bisher manueller Prozessschritte durch den Einsatz technisch fortgeschrittener Werkzeuge wie Business Intelligence-Systeme von vielen angestrebt.

Mit dem *IT Performance Measurement Maturity Model* (ITPM³) wird ein Reifegradmodell präsentiert, das Entwicklungsstufen in der Gestaltung des IT Performance Measurement in Organisationen identifiziert. Das Reifegradmodell kann als Unterstützung dienen, den Ist-Zustand des IT Performance Measurement in Unternehmen strukturiert zu analysieren und darauf basierend Zielzustände und Verbesserungsmaßnahmen zu

bestimmen. Zielgruppe zur Anwendung dieses Modells sind insbesondere IT-Manager sowie externe Berater.

Das Modell wurde – dem Paradigma der Design Science folgend – in einem iterativen Durchlauf der Phasen Artefaktgestaltung und -evaluation entwickelt. Dieser Forschungsprozess soll in weiteren Arbeiten fortgesetzt werden. Dabei soll insbesondere die empirische Basis ausgeweitet werden. Um Unternehmen zur Bekanntgabe von Daten zu der BI-Systemunterstützung ihres IT Performance Measurement zu veranlassen, wird die Entwicklung eines Internetportals angestrebt, mit dessen Hilfe Unternehmen eine Reifegradermittlung vornehmen können. Die Angabe der Ausprägungen der Modellkriterien, die zu Evaluationszwecken genutzt werden können, soll dabei mit der Erhebung der Akzeptanz des Modells durch die Teilnehmer verbunden werden. Die im Rahmen der Modellentwicklung durchgeführten Interviews mit Experten des IT-Managements legen die Vermutung nahe, dass in der Unternehmenspraxis aufgrund des Interesses an der zu Grunde liegenden Problemstellung eine hinreichend große Bereitschaft zur Teilnahme an einem solchen Forschungsansatz vorhanden sein könnte.

5 Literaturverzeichnis

- [BGK04] Bendl, H.; Gleich, R.; Kraus, P.: Wettbewerbsvorteile durch strategieorientierte Steuerung der IT. HMD: Praxis der Wirtschaftsinformatik, 41 (239) 2004; S. 39-47.
- [CG04] Chamoni, P.; Gluchowski, P.: Integrationstrends bei Business-Intelligence-Systemen. Wirtschaftsinformatik, 46 (2) 2004; S. 119-128.
- [CM06] CMMI Product Team: CMMI for Development, 2006. <http://www.sei.cmu.edu/pub/documents/06.reports/pdf/06tr008.pdf>. Abruf am 16.03.2007.
- [Ec06] Eckerson, W.: Performance Dashboards: Measuring, Monitoring, and Managing Your Business. J. Wiley, Hoboken, NJ, 2006.
- [EHH06] Eul, M.; Hanssen, S.; Herzwurm, G.: Systematische Leistungsbestimmung in der IT. Controlling (1) 2006; S. 25-30.
- [Ge02] Gerick, T.: IT-Controlling – Auf dem Weg zum informationstechnischen Wissenszentrum. ControllerNews (2 Sonderdruck) 2002; S. 2-4.
- [GG00] Grothe, M.; Gentsch, P.: Business Intelligence: aus Informationen Wettbewerbsvorteile gewinnen. Addison-Wesley, München, 2000.
- [Gl02] Gladen, W.: Performance Measurement als Methode der Unternehmenssteuerung. HMD: Praxis der Wirtschaftsinformatik, 39 (227) 2002; S. 5-16.
- [GS01] Van Grembergen, W.; Saull, R.: Aligning Business and Information Technology through the Balanced Scorecard at a Major Canadian Financial Group. Its Status Measured with an IT BSC Maturity Model. In: Proceedings of the 34th International Conference on System Sciences, Hawaii, 2001.
- [He04] Hevner, A.R.; March, S.T.; Park, J.; Ram, S.: Design Science in Information Systems Research. MIS Quarterly, 28 (1) 2004; S. 75-105.
- [HL05] Heinrich, L.J.; Lehner, F.: Informationsmanagement: Planung, Überwachung und Steuerung der Informationsinfrastruktur. Oldenbourg, München, 2005.

- [HV93] Henderson, J.C.; Venkatraman, N.: Strategic Alignment: Leveraging Information technology for transforming organizations. IBM Systems Journal, 38 (2&3) 1993; S. 472-484.
- [ITG05] IT Governance Institute: CobiT 4.0. The IT Governance Institute, 2005.
- [Kn06] Kneuper, R.: CMMI: Verbesserung von Softwareprozessen mit Capability Maturity Model Integration. dpunkt-Verlag, Heidelberg, 2006.
- [LWS06] Lehner, F.; Wildner, S.; Scholz, M.: Wirtschaftsinformatik. Eine Einführung. Hanser, München, 2006.
- [MTS06] Müller, A.; von Thiemen, L.; Schröder, H.: IT-Controlling. So messen Sie den Beitrag der Informationstechnologie zum Unternehmenserfolg. Der Controlling-Berater (1) 2006; S. 99-122.
- [No79] Nolan, R.L.: Managing the Crisis in data Processing. Harvard Business Review, 57 (2) 1979; S. 115-126.
- [Pa93] Paulk, M.; Curtis, B.; Chrissis, M.; Weber, C.: Capability Maturity Model for Software, Version 1.1, 1993. <http://www.sei.cmu.edu/pub/documents/93.reports/pdf/tr24.93.pdf>. Abruf am 05.03.2008.
- [Ra04] Rau, K.G.: The CIO Dashboard Performance Management Program: Measuring and Managing the Value of IT. Information Strategy: The Executive's Journal, Winter 2004; S. 6-17.
- [Re04] Renken, J.: Developing an IS/ICT Management Capability Maturity Framework. In: Proceedings of the Research Conference of the South African Institute for Computer Scientists and Information Technologists (SAICSIT), Stellenbosch, 2004; S. 53-62.
- [Th04] Thamm, J.: Assessments von IT-Organisationen. Information Management & Consulting, 19 (1) 2004; S. 65-71.
- [WK02] Wettstein, T.; Küng, P.: A Maturity Model for Performance Measurement Systems, Department of Informatics, Fribourg University, Schweiz, 2002.
- [ZB04] Zarnekow, R.; Brenner, W.: Integriertes Informationsmanagement: Vom Plan, Build, Run zum Source, Make, Deliver. In Zarnekow, R.; Brenner, W.; Grohmann, H.H., (Hrsg.): Informationsmanagement: Konzepte und Strategien für die Praxis., Heidelberg, 2004; S. 3-24.

6 Anhang

Interviewleitfaden

Rahmendaten (liegen bereits vor dem Interview vor)

- Name des Gesprächspartners.
- Position des Gesprächspartners.
- Unternehmen bei dem der Gesprächspartner tätig ist.
- Profil des Unternehmens.

- Kontaktdaten des Gesprächspartners.
- Datum und Uhrzeit des Gesprächstermins.
- Durchführungsart (telefonisch oder vor Ort) und ggf. Ort.

Einleitung

- Vorstellung des Interviewers.
- Dank für die Bereitschaft zur Teilnahme.
- Ziel des Interviews.
- Hinweis auf geplante Studie.
- Klärung formaler Fragen (Erlaubnis zum Mitschneiden des Gesprächs, Anonymisierung auf Wunsch, usw.).
- Kurze Beschreibung des Ablaufs.
- Klärung offener Fragen.

Fragen zum Gesprächspartner

- Persönlicher Werdegang des Gesprächspartners?
- Aktueller Tätigkeits- bzw. Verantwortungsbereich und Aufgabenprofil?

Fragen zur IT-Organisation des Unternehmens

- Interner Aufbau der IT-Organisation?
- Einordnung und Stellenwert der IT-Organisation im Unternehmen?
- Interne und externe Kunden?
- Bedeutung von Outsourcing?

Informationsversorgung des Gesprächspartners

- Informationsbedarf des Gesprächspartners?
- Aktuelle Informationsversorgung des Gesprächspartners?
- Rolle von Berichten?
- Rolle von Tabellenkalkulationen?
- Einsatz von Anwendungssystemen zur Informationsbereitstellung?
- Informationsbereitstellung zufriedenstellend?
- Ansprüche an die Darstellung von Informationen?

IT Performance Measurement

- Erhebung von Daten über die Informationsinfrastruktur?
- Kosten- und Leistungsrechnungsrechnung von IT-Leistungen?
- Einsatz von computergestützten Werkzeugen zur Erhebung?
- Zugriff auf erhobene Daten?
- Einsatz von Business Intelligence-Systemen?
- Verwendung von Kennzahlen und Kennzahlensystemen?
- Anwendungsbereich der Kennzahlen?
- Rolle von prozessorientierten Referenzmodellen wie CobiT oder ITIL?
- IT-Strategie und deren Abstimmung mit der Geschäftsstrategie?

Überprüfung der ersten Version des Reifegradmodells

- Kriterien zur Beurteilung der Qualität bzw. Reife einer computerbasierten Werkzeugunterstützung für das IT-Management?
- Ideale computerbasierte Werkzeugunterstützung für das IT-Management?
- Vorstellung der ersten Version des Reifegradmodells mit seinen vier Stufen.
- Plausibilität des Reifegradmodells und der Stufen?
- Kritik und Anregungen?
- Diskussion.

Abschluss

- Klärung offener Fragen.
- Weiteres Interesse an diesem Thema und einer anschließenden Studie?
- Dank für die Teilnahme.
- Verabschiedung

Bereitstellung und Nutzung von Endbenutzungs-Metadaten im Kontext der Informationslogistik: Empirische Erkenntnisse und Gestaltungsempfehlungen

Tobias Bucher, Ulrich Wlk

Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedberg-Strasse 8
CH-9000 St. Gallen
{tobias.bucher | ulrich.wlk}@unisg.ch

Abstract: Metadaten dienen der Anpassung und wechselseitigen Abstimmung derjenigen Einzelkomponenten, die das System der integrierten Informationslogistik konstituieren. Diese Anpassungs- und Abstimmungsprozesse finden im Rahmen von Planung, Entwicklung, Betrieb und Nutzung der Informationslogistik statt. Der vorliegende Beitrag fokussiert auf den Nutzungsaspekt und damit auf die Betrachtung von sog. Endbenutzungs-Metadaten. Als Endbenutzungs-Metadaten werden solche Metadaten bezeichnet, die darauf abzielen, der Fachseite einen inhaltlichen Überblick über die im Rahmen der Informationslogistik bereitgestellten Daten zu vermitteln und unter den Mitarbeitenden ein einheitliches Verständnis bezüglich der verwendeten Terminologie, der Qualität und der Herkunft der Daten zu schaffen. Der Beitrag skizziert die Ergebnisse einer früheren empirischen Untersuchung und präsentiert Resultate einer Wiederholung ebendieser explorativen Studie. Für die Wiederholung wurden 18 Expertinnen und Experten aus sieben Unternehmen befragt. Aufbauend auf dieser empirischen Grundlage werden sodann Empfehlungen für die Ausgestaltung der Bereitstellung und Nutzung von Endbenutzungs-Metadaten präsentiert. Diese Empfehlungen wurden im Rahmen eines Workshops mit Praxisvertreterinnen und -vertretern erarbeitet und stellen eine wertvolle Gestaltungshilfe für das Metadaten-Management dar. Dabei erhebt der Beitrag aufgrund des Forschungsvorgehens keinen Anspruch auf uneingeschränkte Verallgemeinerbarkeit der Ergebnisse. Vielmehr hat der Beitrag zum Ziel, Anwenderinnen und Anwendern im Umfeld der Informationslogistik interessante und hilfreiche Erkenntnisse zu vermitteln und Empfehlungen zu geben, die auf der langjährigen Erfahrung der befragten Expertinnen und Experten gründen.

1 Metadaten im Kontext der Informationslogistik

Informationslogistik, verstanden als „Planung, Steuerung, Durchführung und Kontrolle der Gesamtheit der Datenflüsse [...], die über eine Betrachtungseinheit hinausgehen, sowie [...] Speicherung und Aufbereitung dieser Daten“ [Wi08, S. 2], zielt auf die Versorgung der Mitarbeitenden eines Unternehmens mit handlungs- und entscheidungsrelevanten Daten ab. Der Betrieb und die Nutzung der Informationslogistik erfordern neben

verschiedenartigen Hard- und Softwarekomponenten auch eine Vielzahl an unterstützenden Prozessen.¹ Zu diesen Prozessen zählt neben dem Architekturmanagement, dem Qualitätsmanagement, dem Management von Datenschutz und Datensicherheit sowie der Gestaltung der Aufbau- und Ablauforganisation auch das Metadatenmanagement [Wi08, S. 10].

Integrierte Informationslogistik wird dementsprechend durch eine Vielzahl an physischen wie auch nicht-gegenständlichen Einzelkomponenten konstituiert. Diese Einzelkomponenten müssen so gut als möglich angepasst und aufeinander abgestimmt werden, um die Nutzeneffekte der Informationslogistik effizient realisieren zu können. Diese Abstimmungsprozesse können unter anderem durch eine integrierte Beschreibung der Einzelkomponenten erleichtert werden [Sc00, S. 102-103]. Solche Beschreibungen werden im Allgemeinen als „Metadaten“ bezeichnet. Dieses Verständnis wird auch in der von ROWOHL ET AL. geprägten Definition des Metadaten-Begriffs für den Kontext der Informationslogistik offenbar:

„Metadaten im Rahmen des Data Warehousing sind all diejenigen Daten, die geeignet sind, Fragen über die im Data Warehouse enthaltenen Daten, deren Transformation und des sie umgebenden Systems zu beantworten.“ [RSS00, S. 5]

Die Anpassungs- und Abstimmungsprozesse zwischen den Einzelkomponenten der integrierten Informationslogistik finden im Rahmen aller Stufen des sog. Informationslogistik-Lebenszyklus, d.h. im Rahmen von Planung, Entwicklung, Betrieb und Nutzung², statt. Diese Unterscheidung findet sich sinngemäss auch in der von MELCHERT vorgeschlagenen Definition des Begriffs „Metadaten“:

„Data-Warehouse-Metadaten abstrahieren im Gegensatz zu DWH-Daten von den Sachverhalten des in einem Data Warehouse (DWH) abgebildeten Realweltausschnitts. Sie beziehen sich auf die im DWH abgelegten Daten sowie auf die damit verbundenen Informationsverarbeitungsprozesse, Informationssystemkomponenten und die organisatorischen Aspekte des Data Warehousing. Im Rahmen von DWH-Entwicklung und DWH-Betrieb dienen Metadaten primär der technischen Dokumentation und/oder Steuerung des DWH-Systems, während ihnen im Rahmen der DWH-Nutzung vor allem die Aufgabe zukommt, den fachlichen und datenverarbeitungsbezogenen Kontext der DWH-Daten zu dokumentieren.“ [Me06, S. 24]

¹ Um den betrachtungseinheitenübergreifenden, auf die Realisierung von Synergien abzielenden Charakter der Informationslogistik zu betonen, bezeichnen WINTER ET AL. [Wi08, S. 9] die Gesamtheit dieser unterstützenden Prozesse als „Infrastruktur“. Die einzelnen Prozesse bzw. Aspekte werden als „Infrastrukturthemen“ oder „Querschnittsthemen“ bezeichnet. Aufgrund fehlender Eindeutigkeit hinsichtlich Definition, Verständnis und Verwendung des Begriffs „Infrastruktur“ findet dieser im vorliegenden Beitrag keine Verwendung.

² KLESSE unterscheidet zwischen den „Teilprozesse[n] der Planung, der Entwicklung, des Betriebs und der analytischen Nutzung eines DWH-Informationssystems“ [KI07, S. 29].

Analog zu diesem Verständnis findet sich in der Literatur eine Vielzahl von Vorschlägen zur Kategorisierung von Metadaten, welche zumeist auf den Verwendungszweck der Metadaten abstellen. So unterscheidet bspw. TOZER zwischen sog. „betrieblichen Metadaten“ (im Original „operational metadata“) einerseits und sog. „informativischen Metadaten“ (im Original „informational metadata“) andererseits [To99, S. 120-125]. Metadaten der erstgenannten Kategorie unterstützen primär Planung, Entwicklung und Betrieb der Informationslogistik, während Metadaten der letztgenannten Kategorie im Wesentlichen an die Fachseite, d.h. an diejenigen Mitarbeitenden, die die Produkte und Leistungen der Informationslogistik nutzen, gerichtet sind. Die informativischen Metadaten zielen demzufolge darauf ab, den Fachbereichsmitarbeitenden einen inhaltlichen Überblick über die bereitgestellten Daten zu vermitteln und ein einheitliches Verständnis bezüglich der verwendeten Terminologie, der Qualität und der Herkunft der angebotenen Daten zu schaffen.

Viele Autorinnen und Autoren, so z.B. AUTH [Au03, S. 44-47], HAYNES [Ha04, S. 15-18] und FOSHAY ET AL. [FMT07, S. 72], schlagen feiner granulare, die von TOZER verwendeten Oberkategorien detaillierende Systematisierungen für Metadaten vor. Dabei wird stets betont, dass diese Systematisierungen per se nicht vollkommen trennscharf sein können. So lassen sich einzelne Metadaten-Elemente durchaus mehreren Kategorien zurechnen, und einzelne diese Kategorien wiederum lassen sich unter Umständen nicht ohne stark vereinfachende oder gar einschränkende Annahmen in eine der beiden von TOZER vorgeschlagenen Oberkategorien einordnen.³

Der vorliegende Beitrag fokussiert auf Metadaten, die im Rahmen der Nutzung der Informationslogistik zum Einsatz kommen, d.h. auf die Oberkategorie der sog. „informativischen Metadaten“. Aufgrund ihres Verwendungszwecks und ihrer expliziten Fokussierung auf Fachbereichsmitarbeitende, die Produkte und Leistungen der integrierten Informationslogistik konsumieren, werden diese Metadaten im Folgenden als „Endbenutzungs-Metadaten“ bezeichnet.

Auf die Wichtigkeit und Bedeutung von Endbenutzungs-Metadaten im Kontext der Informationslogistik wird in zahlreichen wissenschaftlichen wie auch praxisorientierten Beiträgen und Untersuchungen verwiesen. So betont bspw. GLEASON [Gl97, S. 148-149] die Relevanz einer möglichst einfachen und für die Fachseite verständlichen Darstellung von und Versorgung mit Metadaten, die für eine effektive Nutzung der Informationslogistik unerlässlich sind. Die Hypothese, dass wahrgenommene Qualität und Nützlichkeit von Endbenutzungs-Metadaten letztlich entscheidend für Erfolg oder Misserfolg der

³ Als Beispiel hierfür seien Metadaten-Elemente genannt, die Auskunft über Datenquellen und vorgenommene Transformationen der Daten auf dem Weg von den Quellsystemen in eine zentrale Datenbasis (bspw. das Data Warehouse) geben. Solche Metadaten-Elemente fallen, der Systematisierung von AUTH [Au03, S. 44-47] folgend, in die Kategorien „Datentransformation“ und „Systembezug“. FOSHAY ET AL. [FMT07, S. 72] bezeichnen die entsprechende Kategorie als „Transformation“ (im Original „lineage“). Metadaten-Elemente dieser Kategorien können, der Kategorisierung von TOZER [To99, S. 120-125] folgend, wiederum sowohl als „betriebliche Metadaten“ als auch als „informativische Metadaten“ im Rahmen von Planung, Entwicklung, Betrieb und Nutzung des Systems „Informationslogistik“ Verwendung finden.

Informationslogistik sein können, wird von FOSHAY ET AL. [FMT07, S. 73-75] aufgegriffen und empirisch validiert.

Der vorliegende Beitrag hat zum Ziel, die Ergebnisse einer explorativ ausgerichteten empirischen Untersuchung bezüglich der Bereitstellung und Nutzung von Endbenutzungs-Metadaten zu präsentieren und mit den Resultaten einer früheren Untersuchung zu vergleichen. Darauf aufbauend werden sodann Empfehlungen für die Gestaltung des Managements von Metadaten, die sich primär an Fachbereichsmitarbeitende richten, abgeleitet. Diese Empfehlungen beziehen sich in Teilen auf die Schlussfolgerungen der erwähnten früheren Untersuchung, erweitern diese jedoch signifikant in Umfang und Aussagekraft und stellen die wesentlichen Erkenntnisse der beiden empirischen Untersuchungen in pointierter Form dar. Dementsprechend gliedert sich der Beitrag wie folgt: Im nachfolgenden Abschnitt 2 werden die wesentlichen Ergebnisse der den Ausgangspunkt des vorliegenden Beitrags darstellenden Untersuchung von FOSHAY ET AL. [FMT07] skizziert. In Abschnitt 3 werden Vorgehen und Ergebnisse einer Wiederholung ebendieser Untersuchung mit 18 Expertinnen und Experten aus sieben Unternehmen dargestellt. Aufbauend auf einem Vergleich der Ergebnisse der beiden Untersuchungen stellt Abschnitt 4 detaillierte Gestaltungsempfehlungen dar. Diese Empfehlungen wurden im Rahmen eines Workshops mit Praxisvertreterinnen und -vertretern erarbeitet. Abschnitt 5 fasst die wesentlichen Aussagen des vorliegenden Beitrags zusammen und diskutiert mögliche Einschränkungen bezüglich der Allgemeingültigkeit und des Empfehlungscharakters der Erkenntnisse und der Folgerungen.

2 Studie zur Bereitstellung und Nutzung von Endbenutzungs-Metadaten

Wesentliche Grundlage für den vorliegenden Beitrag stellt eine Forschungsarbeit von FOSHAY ET AL. dar, welche im Jahr 2007 publiziert wurde [FMT07]. Die Arbeit geht der Frage nach, ob Endbenutzungs-Metadaten im Kontext der Informationslogistik einen signifikanten Wertbeitrag leisten können, und verfolgt vier Ziele [FMT07, S. 73]:

- Vorschlag einer Kategorisierung von Endbenutzungs-Metadaten
- Erkenntnisgewinn bezüglich der Metadaten-Elemente, die der Fachseite typischerweise zur Verfügung gestellt werden
- Erkenntnisgewinn bezüglich der wahrgenommenen Nützlichkeit von und Zufriedenheit mit der Bereitstellung der Endbenutzungs-Metadaten
- Validierung eines Kausalmodells, das die vier Aspekte „Wahrgenommene Qualität der Endbenutzungs-Metadaten“, „Einstellung der Fachseite zum Data Warehouse“, „Wahrgenommene Nützlichkeit und Bedienbarkeit des Data Warehouse“ sowie „Tatsächliche Nutzung des Data Warehouse“ miteinander in Verbindung setzt

Bezüglich der Kategorisierung von Endbenutzungs-Metadaten differenzieren FOSHAY ET AL. zwischen den vier Metadaten-Klassen „Definition“ (im Original „definitional“),

„Qualität“ (im Original „data quality“), „Navigation“ (im Original „navigational“) und „Transformation“ (im Original „lineage“) [FMT07, S. 72]. Eine Erläuterung dieser Klassen sowie einschlägige Leitfragen sind in Abbildung 1 dargestellt.

Kategorie	Erläuterung
Definition	Metadaten der Kategorie „Definition“ umfassen Elemente, die der Fachseite Auskunft über die fachliche Bedeutung von Daten in den DWH- bzw. BI-Systemen geben. Dies umfasst bspw. die Definition der verwendeten Begriffe bzw. Begriffssysteme, von Berechnungsregeln oder von zulässigen Werten sowie die Angabe von Inhaltsbeispielen. Metadaten dieser Kategorie beantworten folgende Frage: „ Welche fachliche Bedeutung haben die Daten in den DWH- bzw. BI-Systemen? “
Qualität	Metadaten der Kategorie „Qualität“ umfassen Elemente, die der Fachseite Auskunft über die Aktualität, Genauigkeit, Gültigkeit und Vollständigkeit von Daten in den DWH- bzw. BI-Systemen geben. Metadaten dieser Kategorie beantworten folgende Frage: „ Besitzen die Daten in den DWH- bzw. BI-Systemen eine hinreichende Qualität, um sie für einen bestimmten fachlichen Anwendungszweck zu nutzen? “
Navigation	Metadaten der Kategorie „Navigation“ versetzen die Mitarbeitenden der Fachbereiche in die Lage, in DWH- bzw. BI-Systemen nach vorhandenen Daten und deren wechselseitigen Abhängigkeiten zu suchen. Metadaten dieser Kategorie beantworten folgende Frage: „ Wo sind die Daten zu finden, die für einen bestimmten fachlichen Anwendungszweck benötigt werden? “
Transformation	Metadaten der Kategorie „Transformation“ umfassen Elemente, die den Fachbereichsmitarbeitenden Auskunft über die ursprünglichen Quellen der Daten in den DWH- bzw. BI-Systemen geben. Des Weiteren beschreiben diese Metadaten, welche Manipulationen (bspw. Bereinigung, Transformation oder Aggregation) vorgenommen wurden. Metadaten dieser Kategorie beantworten folgende Frage: „ Woher stammen die Daten in den DWH- bzw. BI-Systemen, und welche Manipulationen wurden durchgeführt? “

Abbildung 1: Kategorisierung von Endbenutzungs-Metadaten nach FOSHAY ET AL. [FMT07, S. 72]

Zur Adressierung der drei letztgenannten Forschungsziele führten FOSHAY ET AL. eine empirische Studie durch. Zunächst wurden Vertreterinnen und Vertreter von 266 Unternehmen explorativ bezüglich der Metadaten-Elemente befragt, die den Fachbereichen zur Verfügung gestellt werden (vgl. Abbildung 2). Der o.g. Kategorisierung folgend versorgen mehr als drei Viertel der befragten Unternehmen die Fachseite mit Metadaten-Elementen der Kategorie „Definition“. Metadaten-Elemente der Kategorie „Qualität“ werden in etwas mehr als der Hälfte der Unternehmen angeboten. Die Kategorien „Navigation“ und „Transformation“ werden den Fachbereichsmitarbeitenden nur in knapp bzw. gut einem Drittel der befragten Unternehmen bereitgestellt.

Definition	Qualität	Navigation	Transformation
77.7% der Unternehmen	51.0% der Unternehmen	31.9% der Unternehmen	36.3% der Unternehmen

Abbildung 2: Bereitstellung von Endbenutzungs-Metadaten [FMT07, S. 75]

Im Anschluss daran wurden 268 Angehörige primär technisch orientierter Anspruchsgruppen (d.h. Personen, die für Planung, Entwicklung und Betrieb der Informationslogistik zuständig sind) und 621 Angehörige primär fachlich orientierter Anspruchsgruppen (d.h. Personen, die als Fachbereichsmitarbeitende die Produkte und Leistungen der In-

formationslogistik konsumieren) ebenfalls explorativ bezüglich wahrgenommener Nützlichkeit von und wahrgenommener Zufriedenheit mit den bereitgestellten Endbenutzungs-Metadaten befragt.⁴ Die Angehörigen technisch orientierter Anspruchsgruppen sollten dabei die Perspektive der Fachseite einnehmen, d.h. die Nützlichkeit für und die Zufriedenheit der Fachbereichsmitarbeitenden einschätzen. Im Ergebnis zeigen sich signifikante Unterschiede zwischen den relativen Einschätzungen der technisch orientierten Anspruchsgruppen einerseits und der fachlich orientierten Anspruchsgruppen andererseits.

Fachbereichsmitarbeitende empfinden Metadaten der Kategorien „Qualität“ und „Navigation“ im Vergleich zu den Kategorien „Definition“ und „Transformation“ als deutlich nützlicher. Angehörige technisch orientierter Anspruchsgruppen sind hingegen der Meinung, dass Metadaten der Kategorie „Definition“ die grössten Nutzeneffekte im Zuge der Nutzung der Produkte und Leistungen der Informationslogistik entfalten würden. Den Kategorien „Qualität“, „Navigation“ und „Transformation“ wird nachrangige Nützlichkeit beigemessen (vgl. Abbildung 3).

Anspruchsgruppe	Definition	Qualität	Navigation	Transformation
Primär technisch	8.0	7.5	6.9	6.6
Primär fachlich	6.7	7.5	7.2	6.7

Abbildung 3: Wahrgenommene Nützlichkeit von Endbenutzungs-Metadaten [FMT07, S. 74]

Ein ähnliches Bild ergibt sich auch in Bezug auf die wahrgenommene Zufriedenheit. Fachbereichsmitarbeitende sind vor allem mit Metadaten der Kategorien „Qualität“, „Definition“ und „Navigation“ zufrieden. Technisch orientierte Anspruchsgruppen wiederum schätzen, dass die Fachseite die höchste Zufriedenheit im Zusammenhang mit Metadaten der Kategorien „Definition“ und „Qualität“ empfinden würde (vgl. Abbildung 4).

Anspruchsgruppe	Definition	Qualität	Navigation	Transformation
Primär technisch	5.9	5.5	5.1	5.1
Primär fachlich	5.4	6.2	5.5	4.8

Abbildung 4: Wahrgenommene Zufriedenheit mit Endbenutzungs-Metadaten [FMT07, S. 74]

Zuletzt befragten FOSHAY ET AL. die vorgenannten Personengruppen nach ihrer Einschätzung bezüglich sieben Faktoren, die gemeinsam ein Kausalmodell konstituieren. Der konfirmatorischen Analyse liegen drei Hypothesen zugrunde [FMT07, S. 74]:

- Hypothese H1: Die Qualität und die Nutzung von Endbenutzungs-Metadaten beeinflussen die Einstellung der Fachseite zum Data Warehouse.

⁴ Den Angaben zu Nützlichkeit und Zufriedenheit liegt eine neunstufige Ratingskala zugrunde, wobei die Extremausprägungen 1 für „gar nicht nützlich“ bzw. „sehr unzufrieden“ und 9 für „sehr nützlich“ bzw. „sehr zufrieden“ stehen.

- Hypothese H2: Die Einstellung der Fachseite zum Data Warehouse beeinflusst sowohl die subjektiv wahrgenommene Nützlichkeit als auch die subjektiv empfundene Bedienbarkeit des Data Warehouse.
- Hypothese H3: Die Wahrnehmungen bezüglich der Nützlichkeit und der Bedienbarkeit des Data Warehouse beeinflussen schlussendlich die tatsächliche Nutzung des Data Warehouse.

Für alle vorgenannten Hypothesen werden zudem positive Zusammenhänge unterstellt, d.h. dass eine Zunahme in der bzw. den Eingangsgrößen eine Zunahme in der bzw. den Ausgangsgrößen bedingt. Im Zuge der empirischen Validierung des Kausalmodells konnten alle drei Hypothesen bestätigt werden. Es bestehen demnach statistisch signifikante Kausalzusammenhänge zwischen den vier Aspekten „Wahrgenommene Qualität der Endbenutzungs-Metadaten“, „Einstellung der Fachseite zum Data Warehouse“, „Wahrgenommene Nützlichkeit und Bedienbarkeit des Data Warehouse“ sowie „Tatsächliche Nutzung des Data Warehouse“ [FMT07, S. 74-75].

3 Wiederholung der explorativen Untersuchung und Vergleich der Ergebnisse

Im Zuge der Aufbereitung des vorliegenden Beitrags wurde der explorative Teil der Untersuchung von FOSHAY ET AL. mit 18 Vertreterinnen und Vertretern aus sieben Unternehmen wiederholt. Dabei sollte der Frage nachgegangen werden, ob die – zum Teil erstaunlichen und kontraintuitiven – Ergebnisse von FOSHAY ET AL. reproduziert werden können.⁵

Die befragten Personen arbeiten für Unternehmen, die in den Branchen Finanzdienstleistung, Energieversorgung und Maschinenbau tätig sind. Knapp zwei Drittel (61.1%) der befragten Personen stammen aus IT-Bereichen (Fokus auf Informationslogistik), gut ein Drittel (38.9%) ist in verschiedenen Fachbereichen der Unternehmen tätig oder übt Schnittstellenfunktionen zwischen Fachbereichen und IT aus. Die Befragung wurde mit Hilfe des im Anhang dargestellten standardisierten Fragebogens durchgeführt. Zuvor wurde den Teilnehmerinnen und Teilnehmern der Zweck der Untersuchung erläutert. Darüber hinaus bestand die Möglichkeit, Verständnisfragen zu den von FOSHAY ET AL. verwendeten Metadaten-Kategorien zu stellen. Keine der befragten Personen kannte zum Zeitpunkt der Befragung die Ergebnisse der Studie von FOSHAY ET AL. Diese wurden erst nach Abschluss der Untersuchung präsentiert und diskutiert.

⁵ Bezüglich der Diskussion wichtiger Implikationen, die sich aus dem gewählten Forschungsvorgehen sowie insbesondere aus der geringen Fallzahl ergeben, vgl. insbesondere Abschnitt 5. Zweifellos werden sowohl Allgemeingültigkeit als auch Empfehlungscharakter der Ergebnisse durch die geringe Anzahl befragter Personen – insbesondere im Zusammenhang mit quantitativen Forschungsmethoden – beeinflusst. Zu bemerken ist in diesem Zusammenhang jedoch, dass die Wiederholungs-Untersuchung neben quantitativen Aspekten auch qualitative Elemente umfasst. In Abschnitt 3 werden zunächst die quantitativen Ergebnisse dargestellt und mit den Ergebnissen von FOSHAY ET AL. verglichen. Die Ergebnisse der qualitativ ausgerichteten Expertinnen- und Experteninterviews werden in Abschnitt 4 präsentiert.

Bezüglich der in den befragten Unternehmen bereitgestellten Kategorien von Endbenutzungs-Metadaten ergibt sich aus der Wiederholung der explorativen Untersuchung das in Abbildung 5 dargestellte Bild.⁶

Definition	Qualität	Navigation	Transformation
57.1% der Unternehmen	71.4% der Unternehmen	57.1% der Unternehmen	42.9% der Unternehmen

Abbildung 5: Bereitstellung von Endbenutzungs-Metadaten
(Ergebnisse der Wiederholungs-Untersuchung)

Knapp drei Viertel der Unternehmen stellen den Fachabteilungen Metadaten der Kategorie „Qualität“ zur Verfügung. Immerhin jeweils mehr als die Hälfte der befragten Unternehmen stellt der Fachseite Metadaten der Kategorien „Definition“ und „Navigation“ zur Verfügung, während weniger als die Hälfte der Unternehmen Metadaten der Kategorie „Transformation“ anbietet. Diese Zahlen unterscheiden sich deutlich von den Ergebnissen von FOSHAY ET AL. (vgl. Abschnitt 2). Insbesondere erscheint interessant, dass sich die Anteile der Kategorien „Definition“ und „Qualität“ in den beiden Untersuchungen genau gegensätzlich verhalten. Zudem zeigen die Ergebnisse der eigenen Studie, dass im Vergleich zu den Ergebnissen von FOSHAY ET AL. signifikant mehr Unternehmen die Fachseite mit Metadaten der Kategorie „Navigation“ versorgen. Die Anteile bezüglich der Metadaten-Kategorie „Transformation“ unterscheiden sich nur marginal.

Mehr als die Hälfte (57.1%) der bei der Wiederholung der Untersuchung befragten Unternehmen stellt sowohl Metadaten der Kategorie „Definition“ als auch der Kategorie „Qualität“ bereit. Auch bei FOSHAY ET AL. ist dies die häufigste Kombination, die in immerhin 42.6% der befragten Unternehmen anzutreffen ist [FMT07, S. 75]. Ähnlich verhalten sich auch die Ergebnisse bezüglich der Kombination aller vier Kategorien von Endbenutzungs-Metadaten. Die eigene Untersuchung zeigt, dass gut ein Viertel (28.6%) der befragten Unternehmen ihren Fachbereichen Metadaten aller vier Kategorien zur Verfügung stellt. FOSHAY ET AL. weisen in ihrer Untersuchung einen entsprechenden Anteil von 13.9% aus [FMT07, S. 75].

Aufgrund der im Vergleich zu den Ergebnissen von FOSHAY ET AL. fast ausnahmslos signifikant höheren Anteile, die in der Wiederholung der Untersuchung bezüglich der Bereitstellung von Endbenutzungs-Metadaten ermittelt wurden, lässt sich schliessen, dass die sieben befragten Unternehmen insgesamt einen höheren Reifegrad in Bezug auf

⁶ Die genannten Anteile beziehen sich auf eine zentrale, für alle interessierten und berechtigten Mitarbeitenden zugängliche Dokumentation der jeweiligen Metadaten, jedoch nicht auf eine konkrete technische Realisierungsvariante des Metadaten-Zugriffs. So wird bspw. nicht zwischen einer integrierten und kombinierten Form der Metadaten-Bereitstellung zusammen mit den Objektdaten (z.B. in Form von Metadaten-Annotationen in Berichten) einerseits und der zentralen Dokumentation der relevanten Metadaten in einem Metadaten-Katalog (z.B. strukturiert in einem Metadaten-Repository oder semi- bzw. unstrukturiert in einem Textdokument) andererseits unterschieden. Relevant für die Beantwortung der Frage, ob Endbenutzungs-Metadaten angeboten werden, ist einzig und allein die zentrale Verfügbarkeit und Zugänglichkeit dieser Metadaten. Metadaten, die nur dezentral (bspw. lokal in den einzelnen Fachbereichen) gepflegt und vor allem vorgehalten werden, gelten dementsprechend als nicht angeboten bzw. nicht bereitgestellt.

die Bereitstellung von Endbenutzungs-Metadaten aufweisen. Abstrahiert man von Stichprobenfehlern und anderen verzerrenden Effekten⁷, so mag ein möglicher Erklärungsansatz hierfür in der Branchenzugehörigkeit der in die Wiederholung der Untersuchung einbezogenen Unternehmen zu finden sein. Insbesondere Finanzdienstleister und Energieversorger sehen sich bereits seit Langem mit der Herausforderung einer möglichst effizienten Verarbeitung und Analyse von Massendaten konfrontiert. Dementsprechend weisen die in diesen Unternehmen existierenden Informationslogistik-Lösungen im Vergleich zu anderen Branchen häufig höhere Reifegrade auf.

In Bezug auf die wahrgenommene Nützlichkeit von Endbenutzungs-Metadaten ergibt sich aus der Wiederholung der explorativen Untersuchung ein im Vergleich zu den Ergebnissen von FOSHAY ET AL. (vgl. Abschnitt 2) tendenziell ähnliches, in einzelnen Teilen jedoch auch widersprüchliches Bild (vgl. Abbildung 6, Abweichungen im Vergleich zu Abbildung 3 in Klammern dargestellt⁸). Primär technisch orientierte Anspruchsgruppen vertreten die Ansicht, dass Metadaten der Kategorien „Definition“ und „Transformation“ für die Fachseite den höchsten Nutzen entfalten würden. In diesem Punkt stimmen sie mit den Einschätzungen der Fachbereichsmitarbeitenden überein. Gleiches gilt auch für die tendenziell geringere Nützlichkeit der Metadaten-Kategorie „Transformation“. Ein signifikanter, beinahe unerklärbarer Unterschied zwischen den Einschätzungen der technisch orientierten Anspruchsgruppen einerseits und den der Fachbereichsmitarbeitenden andererseits besteht in Bezug auf Endbenutzungs-Metadaten der Kategorie „Qualität“. Technisch orientierte Personen bewerten den Nutzen ebensolcher Metadaten als deutlich höher im Vergleich zu primär fachlich orientierten Anspruchsgruppen. Diese Aussage steht im offensichtlichen Widerspruch zu den Ergebnissen von FOSHAY ET AL., nach denen die Fachseite der Metadaten-Kategorie „Qualität“ höchste Nützlichkeit attestiert (vgl. Abschnitt 2).

Anspruchsgruppe	Definition	Qualität	Navigation	Transformation
Primär technisch	7.4 (-0.6)	6.5 (-1.0)	6.8 (-0.1)	5.4 (-1.2)
Primär fachlich	7.2 (+0.5)	4.7 (-2.8)	6.0 (-1.2)	5.3 (-1.4)

Abbildung 6: Wahrgenommene Nützlichkeit von Endbenutzungs-Metadaten
(Ergebnisse der Wiederholungs-Untersuchung und Abweichungen zur Primär-Untersuchung)

Ein ähnliches Muster ergibt sich auch für die wahrgenommene Zufriedenheit mit den der Fachseite bereitgestellten Endbenutzungs-Metadaten. In der Wiederholung der Untersuchung (vgl. Abbildung 7, Abweichungen im Vergleich zu Abbildung 4 in Klammern

⁷ Beispielsweise wurden die Daten für die Untersuchung von FOSHAY ET AL. im Sommerhalbjahr 2005 gesammelt [FMT07, S. 74], während die vorliegend beschriebene Wiederholung der Befragung im März 2008 stattfand.

⁸ Aufgrund des Forschungsvorgehens (Vergleich zweier empirischer Untersuchungen, wobei bezüglich der Primär-Untersuchung von FOSHAY ET AL. nur auf die veröffentlichten Ergebnisse, nicht jedoch auf den kompletten Datensatz zurückgegriffen werden konnte) ist nur ein mittelwertbezogener Vergleich der wahrgenommenen Nützlichkeit bzw. Zufriedenheit möglich. Aussagen bezüglich der statistischen Signifikanz der Abweichungen können auf dieser Datengrundlage nicht getroffen werden.

dargestellt) bewerten primär technische Anspruchsgruppen die Zufriedenheit der Fachseite insbesondere in Bezug auf die Metadaten-Kategorien „Definition“, „Qualität“ und „Transformation“ als vergleichsweise hoch. In Bezug auf die beiden Kategorien „Definition“ und „Transformation“ stimmen die Einschätzungen der technischen Anspruchsgruppen zumindest tendenziell mit der Wahrnehmung der Fachbereichsmitarbeitenden überein. Hinsichtlich der Kategorie „Qualität“ ist dies jedoch, analog zur vorstehend diskutierten Einschätzung der Nützlichkeit, nicht der Fall. Die Fachseite scheint mit Endbenutzungs-Metadaten der Kategorie „Qualität“ weitestgehend unzufrieden zu sein. Hier ist wiederum ein offensichtlicher Widerspruch zu den Ergebnissen von FOSHAY ET AL. zu beobachten, nach denen die Fachseite mit Metadaten der Kategorie „Qualität“ in hohem Masse zufrieden ist (vgl. Abschnitt 2). Bezüglich der Zufriedenheit mit Endbenutzungs-Metadaten der Kategorie „Navigation“ stimmen technisch und fachlich orientierte Anspruchsgruppen in ihrer Einschätzung der geringsten Nützlichkeit aller vier Kategorien überein.

Anspruchsgruppe	Definition	Qualität	Navigation	Transformation
Primär technisch	6.0 (+0.1)	5.6 (+0.1)	4.4 (-0.7)	5.4 (+0.3)
Primär fachlich	5.2 (-0.2)	4.3 (-1.9)	4.3 (-1.2)	4.7 (-0.1)

Abbildung 7: Wahrgenommene Zufriedenheit mit Endbenutzungs-Metadaten
(Ergebnisse der Wiederholungs-Untersuchung und Abweichungen zur Primär-Untersuchung)

	Definition Nützlichkeit	Definition Zufriedenheit	Qualität Nützlichkeit	Qualität Zufriedenheit	Navigation Nützlichkeit	Navigation Zufriedenheit	Transforma- tion Nützlichkeit	Transforma- tion Zufriedenheit
Definition Nützlichkeit	--	0.170	0.678 **	0.271	0.488	0.332	0.458	0.583
Definition Zufriedenheit	0.170	--	0.367	0.563 *	0.460	0.455	0.420	0.712 *
Qualität Nützlichkeit	0.678 **	0.367	--	0.807 ***	0.438	0.037	0.349	0.443
Qualität Zufriedenheit	0.271	0.563 *	0.807 ***	--	0.553	0.289	0.623	0.648
Navigation Nützlichkeit	0.488	0.460	0.438	0.553	--	0.656 *	0.568	0.758 **
Navigation Zufriedenheit	0.332	0.455	0.037	0.289	0.656 *	--	0.456	0.403
Transforma- tion Nützlichkeit	0.458	0.420	0.349	0.623	0.568	0.456	--	0.891 ***
Transforma- tion Zufriedenheit	0.583	0.712 *	0.443	0.648	0.758 **	0.403	0.891 ***	--

Legende: * Korrelation signifikant ($p < 0.1$), ** Korrelation hoch signifikant ($p < 0.05$), Korrelation höchst signifikant ($p < 0.01$)

Abbildung 8: Pearson-Korrelationsmatrix (Ergebnisse der Wiederholungs-Untersuchung)

Abbildung 8 zeigt die Pearson-Korrelationsmatrix für die Einschätzungen der Nützlichkeit und der Zufriedenheit bezüglich der vier Kategorien von Endbenutzungs-Metadaten. Signifikante Korrelationen (bezogen auf zweiseitige Signifikanztests) sind durch hoch-

gestellte Sterne gekennzeichnet. Alle Korrelationen sind positiv, d.h. eine Erhöhung bzw. Verminderung einer Variablen geht einher mit einer gleichgerichteten Erhöhung bzw. Verminderung der anderen Variable.

Bemerkenswert sind zunächst die signifikanten Zusammenhänge zwischen Nützlichkeit und Zufriedenheit in Bezug auf die Metadaten-Kategorien „Qualität“, „Navigation“ und „Transformation“: Je höher die wahrgenommene Nützlichkeit, desto höher ist auch die Zufriedenheit in Bezug auf derartige Endbenutzungs-Metadaten. Eine Ausnahme von dieser Regel stellt die Kategorie „Definition“ dar. Hier besteht kein signifikanter Zusammenhang zwischen Nützlichkeit und Zufriedenheit. Die Einzelfallanalyse zeigt, dass ein Teil der befragten Personen die Nützlichkeit als sehr hoch, die Zufriedenheit gleichzeitig aber als sehr gering bewertet. Ein anderer Teil schätzt sowohl Nützlichkeit als auch Zufriedenheit als mittelmässig ein.

Des Weiteren sind auch die beiden Abhängigkeiten, die zwischen der wahrgenommenen Nützlichkeit der Kategorien „Definition“ und „Qualität“ sowie zwischen der wahrgenommenen Zufriedenheit mit diesen beiden Metadaten-Kategorien bestehen, auffällig. Endbenutzungs-Metadaten der Kategorien „Definition“ und „Qualität“ weisen dementsprechend sowohl inhaltliche wie auch formale Abhängigkeiten auf. Hohe Nützlichkeit der Endbenutzungs-Metadaten dieser Kategorien gehen ebenso miteinander einher wie hohe Zufriedenheit der Fachseite mit den bereitgestellten definitions- und qualitätsbezogenen Metadaten.

4 Ableitung von Gestaltungsempfehlungen

Im unmittelbaren Anschluss an die Durchführung der vorstehend skizzierten explorativen Untersuchung wurden den Befragten sowohl die Ergebnisse der Forschungsarbeit von FOSHAY ET AL. (vgl. Abschnitt 2) als auch eine Schnellauswertung der Ergebnisse der eigenen Umfrage präsentiert. Nach einer kurzen, eher informellen Einstiegsdiskussion, die primär der Klärung von Verständnisfragen diente, wurden im Rahmen eines Workshops Expertinnen- und Experteninterviews mit den beteiligten Praxisvertreterinnen und -vertretern durchgeführt.⁹ Dabei wurden folgende Leitfragen erörtert:

- Welche konkreten Nutzeneffekte ergeben sich für die Fachseite aus Endbenutzungs-Metadaten der vier Kategorien „Definition“, „Qualität“, „Navigation“ und Transformation“?

⁹ Als Expertinnen- bzw. Experteninterviews werden im Allgemeinen offene Befragungen bezeichnet, in deren Rahmen ein „organisatorischer oder institutioneller Zusammenhang, [...] in dem sie [die Befragten] nur einen ‚Faktor‘ darstellen“ [MN05, S. 72-73] thematisiert wird. Dabei gelten auch vergleichsweise kleine Fallzahlen von ungefähr 20 befragten Personen als zulässig [MN05, S. 72]. Die in die Expertinnen- und Experteninterviews einbezogenen Praxisvertreterinnen und -vertreter verfügen allesamt über langjährige einschlägige Berufserfahrung und besitzen Spezialkenntnisse entweder im Kontext der Planung, der Entwicklung und des Betriebs der Informationslogistik, im Bereich der Nutzung der Informationslogistik, oder in beiden Tätigkeitsfeldern. Dementsprechend können die befragten Personen als Expertinnen und Experten für den Gegenstand und den Kontext der vorliegenden Untersuchung bezeichnet werden (vgl. hierzu bspw. [FI02, S. 139-140; FL05, S. 226-227; MN05, S. 73-75]).

- Welche Implikationen ergeben sich aus dem im Rahmen der explorativen Untersuchungen beobachteten Verhältnis zwischen der Einschätzung der Nützlichkeit von Endbenutzungs-Metadaten einerseits und dem tatsächlichen Angebot sowie der Zufriedenheit der Fachbereichsmitarbeitenden andererseits?
- Welche Priorisierung bzw. zeitliche und logische Reihenfolge bezüglich der Bereitstellung von Metadaten der vier Kategorien „Definition“, „Qualität“, „Navigation“ und „Transformation“ ist sinnvoll, wenn vorausgesetzt wird, dass einerseits alle in diesem Beitrag diskutierten Kategorien von Endbenutzungs-Metadaten relevant und nützlich sind und dass andererseits in fast allen Unternehmen die verfügbaren Ressourcen zur Befriedigung der fachseitigen Metadaten-Bedarfe beschränkt sind?
- Welche verschiedenen Optionen bestehen bezüglich der Ausgestaltung der Bereitstellung von Endbenutzungs-Metadaten, und bestehen diesbezüglich Unterschiede in der Eignung für die vier genannten Metadaten-Kategorien?

Die Diskussionsergebnisse bezüglich der Nutzeneffekte von Endbenutzungs-Metadaten sind in Abbildung 9 überblicksartig zusammengestellt. Im Zuge der Auswertung wurden inhaltlich übereinstimmende bzw. ähnliche Nennungen der Expertinnen und Experten aus der Praxis zusammengefasst.

Kategorie	Nutzeneffekte
Definition	<ul style="list-style-type: none"> ✓ Fördert ein einheitliches Verständnis der Daten ✓ Erleichtert die fachliche Interpretation der Daten und der Auswertungen ✓ Ermöglicht die Erstellung von aussagekräftigen und vergleichbaren Auswertungen ✓ Unterstützt das Anforderungsmanagement in Bezug auf Informationslogistik ✓ Führt mittelbar zu Zeitersparnis und damit zu Kostenreduktion
Qualität	<ul style="list-style-type: none"> ✓ Ermöglicht den Nachweis der Datenintegrität ✓ Erleichtert die Beurteilung der Aktualität von Daten und Auswertungen ✓ Unterstützt die Einschätzung der Vollständigkeit der Daten ✓ Fördert das Vertrauen der Nutzerinnen und Nutzer der Informationslogistik in die Daten ✓ Führt damit mittelbar zu höherer Sicherheit durch die Reduktion von Risiken
Navigation	<ul style="list-style-type: none"> ✓ Erleichtert die Suche nach Daten und Auswertungen ✓ Fördert die Autonomie der Fachbereiche ✓ Ermöglicht die Beurteilung des Kontexts der Daten und der Auswertungen ✓ Erleichtert die Bedienbarkeit der Informationslogistik-Lösungen für die Fachseite ✓ Führt mittelbar zu Zeitersparnis und damit zu Kostenreduktion
Transformation	<ul style="list-style-type: none"> ✓ Ermöglicht Aussagen über Herkunft der Daten und erleichtert so deren Interpretation ✓ Fördert dadurch auch die Nachvollziehbarkeit von Daten und Auswertungen ✓ Unterstützt Vergleichbarkeit und führt zu höherer Transparenz ✓ Erleichtert die Plausibilisierung von evtl. fragwürdigen Resultaten ✓ Führt damit mittelbar zu einer Reduktion potenzieller Fehler und zu Kostenreduktion

Abbildung 9: Nutzeneffekte von Endbenutzungs-Metadaten

Metadaten der Kategorie „Definition“ informieren die Fachbereichsmitarbeitenden über die fachliche Bedeutung der durch Informationslogistik bereitgestellten Daten und fördern dadurch ein einheitliches Verständnis, vor allem im Hinblick auf Begriffsdefinitionen.

nen und Berechnungsregeln. Dadurch wird auch die fachliche Interpretation der Daten erleichtert und die Vergleichbarkeit von Auswertungen sichergestellt. Weiterhin unterstützen Endbenutzungs-Metadaten der Kategorie „Definition“ auch die Erhebung fachlicher Anforderungen für Produkte und Leistungen der Informationslogistik. Mittelbar führen diese Nutzeneffekte zu Zeitersparnis und damit zu Kostenreduktion.

Endbenutzungs-Metadaten der Kategorie „Qualität“ geben der Fachseite Auskunft über die Aktualität, Genauigkeit, Gültigkeit und Vollständigkeit der bereitgestellten Daten. Durch den Nachweis der Datenintegrität und die Information der Fachbereichsmitarbeitenden über mögliche Datenqualitätsprobleme wird das Vertrauen der Nutzerinnen und Nutzer in die bereitgestellten Daten erhöht. Dies führt zu einer erhöhten Bereitschaft der Fachseite zur Nutzung der Informationslogistik bei gleichzeitiger Verminderung operativer Risiken.

Metadaten der Kategorie „Navigation“ fördern die Autonomie der Fachbereiche durch die Bereitstellung eines Datenverzeichnisses, welches Auskunft über die durch die integrierte Informationslogistik bereitgestellten Daten und Auswertungen gibt. Durch einfachere Bedienbarkeit der Informationslogistik-Lösungen und verbesserte Interpretierbarkeit des Kontexts der Daten und der Auswertungen können wiederum Zeit- und Kosteneinsparungen realisiert werden.

Endbenutzungs-Metadaten der Kategorie „Transformation“ erleichtern ebenfalls die fachseitige Interpretation der Daten und Auswertungen. Informationen bezüglich der Herkunft der Daten und der vorgenommenen Transformationsschritte unterstützen Nachvollziehbarkeit, Vergleichbarkeit und Transparenz. Ausserdem erhält die Fachseite durch die Bereitstellung von Metadaten der Kategorie „Transformation“ ein Instrument, um fragwürdige Resultate der Informationslogistik eigenständig und effizient hinterfragen bzw. plausibilisieren zu können. Mittelbar führt dies zur Reduktion potenzieller Fehlerquellen und zu Kosteneinsparungen.

Die von FOSHAY ET AL. aufgestellte These, dass die vorgeschlagene Kategorisierung von Endbenutzungs-Metadaten ein sinnvolles, intuitiv verständliches und nutzenstiftendes Rahmenwerk für die Spezifikation von fachseitigen Metadaten-Bedarfen darstelle [FMT07, S. 76], konnte durch die Wiederholung der explorativen Untersuchung und im Rahmen der Interviews bestätigt werden. Ebenso wurde offensichtlich, dass sich die Wahrnehmungen bezüglich der Nützlichkeit von Endbenutzungs-Metadaten und der Zufriedenheit der Fachseite mit den bereitgestellten Metadaten signifikant in Abhängigkeit davon unterscheiden, ob primär technisch orientierte Anspruchsgruppen oder primär fachlich orientierte Anspruchsgruppen befragt werden. Auch diese Beobachtung deckt sich mit den Ergebnissen von FOSHAY ET AL. [FMT07, S. 76].

Ein auffälliger Unterschied zwischen den Ergebnissen von FOSHAY ET AL. und den Resultaten der vorliegend beschriebenen Wiederholung ebendieser Untersuchung liegt in der relativen Bewertung der Metadaten-Kategorie „Qualität“ in Bezug auf Nützlichkeit und Zufriedenheit. FOSHAY ET AL. charakterisieren die Bereitstellung qualitätsbezogener Metadaten als wichtigsten Einflussfaktor, um die Zufriedenheit der Fachbereichsmitarbeitenden zu fördern und eine hohe fachseitige Nutzung und Akzeptanz der Informati-

onslogistik zu sichern [FMT07, S. 76-77]. Die Expertinnen und Experten, die an der vorliegend dokumentierten Untersuchung beteiligt waren, charakterisieren Metadaten der Kategorie „Qualität“ zwar als nützlich und räumen der Bereitstellung ebensolcher Metadaten für die Fachbereiche auch gewisse Priorität ein¹⁰, sehen die eigentliche Herausforderung und damit den Schlüssel zu einer erfolgreichen Vermarktung und Verankerung der Informationslogistik in ihren Unternehmen jedoch in der Bereitstellung und Nutzung definitorischer Endbenutzungs-Metadaten.

Metadaten der Kategorie „Definition“ geben Auskunft über die fachliche Bedeutung von Daten, die durch Informationslogistik zur fachseitigen Nutzung verfügbar gemacht werden. Dadurch sind diese Metadaten essentiell für eine effektive Nutzung der Informationslogistik. Gleichzeitig besteht auch ein enger inhaltlicher Zusammenhang zwischen Metadaten der Kategorien „Definition“ und „Transformation“, da aus der Angabe von Ursprungs-Datenquellen und Transformationspfaden (transformationsorientierte Metadaten) häufig – meist implizit – auf die zugrunde liegenden Definitionen, Begriffssysteme und Berechnungsregeln (definitorische Metadaten) geschlossen werden kann. Letztere Metadaten sind grundsätzlich eher technischer Natur, während erstere rein fachliche Metadaten repräsentieren (vgl. bspw. [Au03, S. 39-40; Ma00, S. 49; Me06, S. 23]). Letztendlich stellen diese beiden Kategorien unterschiedliche Sichten auf ein und denselben logischen Metadatenbestand dar.

Die Erhebung und Abstimmung fachlich-definitorischer Metadaten ist jedoch mit vielen grundsätzlichen Herausforderungen verbunden, die im vorliegenden Beitrag nicht erschöpfend diskutiert werden können. Zu diesen Herausforderungen zählt vor allem das in fast allen Unternehmen anzutreffende Phänomen der Begriffsvielfalt, welches wiederum zu zunehmender terminologischer Inkonsistenz führt. Als Beispiel hierfür sei genannt, dass in verschiedenen Fachbereichen eines Unternehmens häufig unterschiedliche Benennungen von Begriffen entstehen, was wiederum die Kommunikation zwischen diesen Bereichen signifikant erschweren oder gar behindern kann. Da Informationslogistik aber gerade darauf abzielt, Synergien durch die übergreifende Nutzung von Daten zu realisieren [Wi08, S. 8-9], besteht eine wesentliche Aufgabe – auch und gerade für das Management von Endbenutzungs-Metadaten – darin, die unkontrollierte Entwicklung von Terminologien zu vermeiden. Deshalb ist es zwingend erforderlich, im Rahmen des Metadatenmanagements auch Terminologiemanagement zu betreiben, d.h. die „Planung, Steuerung und Organisation der in einer spezifischen Bedeutung kontrolliert verwendeten Wörter und Fachausdrücke eines Fachgebiets“ [He97, S. 43]. Mit der Abstimmung und Vereinheitlichung von Terminologien als notwendige Voraussetzung für die Bereitstellung von Endbenutzungs-Metadaten der Kategorie „Definition“ ist ein hoher Mehrwert für die Fachseite verbunden. Mittelfristig führen diese Massnahmen zu einer ver-

¹⁰ Die an den Interviews beteiligten Expertinnen und Experten vertreten die Ansicht, dass qualitätsbezogene Metadaten der Fachseite vergleichsweise einfach und ohne grösseren Aufwand bereitgestellt werden können. Oft genüge hierfür ein einfacher periodischer Qualitätsbericht oder ein entsprechend angebrachter Kommentar, der für die durch Informationslogistik bereitgestellten Auswertungen Auskunft über die Aktualität und die Integrität der zugrunde liegenden Daten (d.h. bezüglich der Übereinstimmung mit den in den Quellsystemen geführten Daten) gäbe. Viele kommerzielle DWH-/BI-Produkte unterstützen derartige Funktionalitäten.

stärkten fachseitigen Nutzung der Informationslogistik und gleichzeitig auch zu einer erhöhten Zufriedenheit der Fachbereichsmitarbeitenden mit den angebotenen Produkten und Leistungen.

Metadaten der Kategorie „Navigation“ erleichtern der Fachseite das eigenständige Auffinden von Daten und Auswertungen, die durch die Informationslogistik-Lösungen bereitgestellt werden. Gleichwohl vertreten die für den vorliegenden Beitrag befragten Expertinnen und Experten die Ansicht, dass dem Aufbau von Datenverzeichnissen – im Unterschied zur vorstehend diskutierten Bereitstellung definitorischer und transformationsbezogener Metadaten – keine verstärkte Aufmerksamkeit geschenkt werden müsse. Vielmehr sollten die Anstrengungen auf die Kategorien „Definition“, „Transformation“ und allenfalls „Qualität“ konzentriert werden. Metadaten der Kategorie „Navigation“ entstünden indessen als Nebenprodukt dieser Bemühungen. Vielfach wird die Erstellung von Datenverzeichnissen und -katalogen auch durch die verwendeten DWH-/BI-Tools unterstützt.

Zusammenfassend lassen sich folgende Gestaltungsempfehlungen ableiten:

- Endbenutzungs-Metadaten der Kategorie „Definition“ sind von übergeordneter Relevanz und Nützlichkeit für die Fachbereichsmitarbeitenden. Deshalb ist ihrer Aufbereitung und Bereitstellung hohe Priorität einzuräumen. Insbesondere ist in diesem Zusammenhang auch auf die bereichsübergreifende Abstimmung von Terminologien zu achten. Endbenutzungs-Metadaten der Kategorie „Transformation“ stehen in engem Zusammenhang zu definitorischen Metadaten. Metadaten dieser beiden Kategorien sollten zentral zusammengeführt, abgestimmt und allen berechtigten Mitarbeitenden zur Verfügung gestellt werden. Die konkrete technische Realisierung ist dabei sekundär. Definitorische Metadaten können eher unstrukturiert in Textdokumenten oder eher strukturiert via dedizierter Metadaten-Repositories bereitgestellt werden. Eine direkte Verlinkung in den Auswertungen enthaltener Objektdaten-Elemente mit den zugehörigen definitorischen Metadaten ist hilfreich.
- Endbenutzungs-Metadaten der Kategorie „Qualität“ sind ebenfalls wichtig und hilfreich für die fachseitige Beurteilung der Aktualität und der Integrität von Daten. Allerdings erfordert die Bereitstellung von qualitätsbezogenen Metadaten vergleichsweise wenig Aufwand. Ein einfacher periodischer Qualitätsbericht oder ein an den Objektdaten-Elementen bzw. den Auswertungen angebrachter Kommentar, der Auskunft über die Aktualität und die Integrität der zugrunde liegenden Daten gibt, ist häufig ausreichend. Viele kommerzielle DWH-/BI-Produkte unterstützen derartige Funktionalitäten.
- Endbenutzungs-Metadaten der Kategorie „Navigation“ kommt allenfalls untergeordnete Bedeutung zu. Sie entstehen häufig als Nebenprodukt von Massnahmen, die auf die Aufbereitung und Bereitstellung von Metadaten der anderen Kategorien abzielen. Vielfach wird die Erstellung navigationsbezogener Metadaten auch durch die verwendeten DWH-/BI-Tools selbst unterstützt, so dass in dieser Beziehung unter Umständen verstärktes Augenmerk auf den Funktionsumfang der eingesetzten Tools zu legen ist.

5 Zusammenfassung und Ausblick

Der vorliegende Beitrag stellt die Ergebnisse zweier empirischer Untersuchungen zur Bereitstellung und Nutzung von sog. Endbenutzungs-Metadaten vergleichend dar und präsentiert darauf aufbauend Gestaltungsempfehlungen für das Metadaten-Management. Als Endbenutzungs-Metadaten werden solche Metadaten bezeichnet, die im Rahmen der Nutzung der Informationslogistik zum Einsatz kommen und sich dabei an Fachbereichsmitarbeitende richten, die Produkte und Leistungen der integrierten Informationslogistik konsumieren.

Die Gestaltungsempfehlungen wurden im Rahmen eines Workshops durch Interviews mit Expertinnen und Experten aus sieben Unternehmen der Branchen Finanzdienstleistung, Energieversorgung und Maschinenbau erarbeitet. Sowohl der in diesem Beitrag beschriebenen Wiederholung der explorativen Untersuchung als auch den Expertinnen- und Experteninterviews liegen vergleichsweise kleine Fallzahlen von weniger als 20 befragten Personen zugrunde. Daraus ergeben sich wichtige Implikationen für den Gültigkeitsbereich und den Empfehlungscharakter der beschriebenen Forschungsergebnisse.

Allgemeingültigkeit und Empfehlungscharakter werden häufig als Gütekriterien für methodenorientierte Vorgehens- und Gestaltungsempfehlungen genannt [Br03, S. 31-32; Wo06, S. 100-101]. Unter dem Kriterium der Allgemeingültigkeit wird demnach verstanden, dass Empfehlungen von mehreren unterschiedlichen Unternehmen aufgegriffen und nutzenstiftend eingesetzt werden können. Empfehlungscharakter impliziert den Anspruch, dass die in diesem Beitrag dargestellten Erkenntnisse und Folgerungen Vorbildcharakter im Sinne einer Soll-Ausgestaltung des Managements von Endbenutzungs-Metadaten besitzen würden.

Aus der gewählten Forschungsmethodik resultieren gewisse Einschränkungen bezüglich der beschriebenen Gütekriterien. Allgemeingültigkeit kann nicht zwingend unterstellt werden, da den Empfehlungen zum einen eine vergleichsweise geringe Fallzahl befragter Personen bzw. Unternehmen zugrunde liegt. Zum anderen genügt auch das Auswahlverfahren der befragten Personen bzw. Unternehmen nicht den Anforderungen einer Zufallsstichprobe. Allerdings wird in der gestaltungsorientierten Forschung, bspw. im Methoden-Engineering (vgl. bspw. [HBO94]) und auch in der Referenzmodellierung (vgl. z.B. [Br03]), stets betont, dass universelle Anwendbarkeit (und damit Allgemeingültigkeit) von Vorgehens- und Gestaltungsempfehlungen kein erreichbares und sinnvolles Ziel sei. Stattdessen wiesen Vorgehens- und Gestaltungsempfehlungen eine inhärente Situationsspezifität auf.

Übertragen auf den Kontext dieses Beitrags bedeutet dies, dass die dargestellten Gestaltungsempfehlungen insbesondere für solche Unternehmen Gültigkeit und Relevanz besitzen, die mit den an der Ableitung ebendieser Empfehlungen beteiligten Unternehmen vergleichbar sind. Mögliche Kriterien, die für einen solchen situationsbezogenen Vergleich herangezogen werden können, sind die Branche (Finanzdienstleistung, Energieversorgung, Maschinenbau), die Unternehmensgrösse (mittelgrosse bis grosse Unternehmen), das Marktumfeld, in dem die Unternehmen tätig sind (Unternehmen aus dem deutschsprachigen Raum, sowohl national als auch international tätig) und der Reifegrad

der Informationslogistik in den Unternehmen (vergleichsweise hohe Reife, in allen Unternehmen wird Informationslogistik seit vielen Jahren erfolgreich eingesetzt).

Empfehlungscharakter kann den Erkenntnissen und Schlussfolgerungen dieses Beitrags insofern zugesprochen werden, als dass an der Erarbeitung Expertinnen und Experten aus der Praxis beteiligt waren, die allesamt über langjährige einschlägige Berufserfahrung verfügen und die Spezialkenntnisse im Kontext der Planung, der Entwicklung und des Betriebs der Informationslogistik, im Bereich der Nutzung der Informationslogistik, oder in beiden vorgenannten Tätigkeitsfeldern besitzen.

Aus der vorstehenden Diskussion möglicher Einschränkungen ergibt sich in logischer Konsequenz weiterer Forschungsbedarf: Zum einen sollte dem Allgemeingültigkeitsanspruch der Gestaltungsempfehlungen insofern Rechnung getragen werden, als dass Vertreterinnen und Vertreter aus Unternehmen, die sich hinsichtlich ihrer Charakteristika von den an der Erarbeitung der Empfehlung beteiligten Unternehmen unterscheiden, die Aussagen kritisch reflektieren und dahingehend prüfen, wie diese angepasst und verändert werden müssten, um in anderen Unternehmen effektiv eingesetzt werden zu können. Zum anderen sollte der Empfehlungscharakter der Aussagen evaluiert werden, indem Unternehmen, deren Charakteristika mit denen der an der Erarbeitung beteiligten Unternehmen übereinstimmen, die Gestaltungsempfehlungen dahingehend überprüfen, ob diese realistisch umsetzbar und gleichzeitig nutzenstiftend sind. Auf diese Art und Weise können die im vorliegenden Beitrag dargestellten Gestaltungsempfehlungen einer kritischen Evaluation unterzogen werden.

6 Literaturverzeichnis

- [Au03] Auth, G.: Prozessorientierte Organisation des Metadatenmanagements für Data-Warehouse-Systeme. Dissertation, Universität St. Gallen, St. Gallen, 2003.
- [Br03] vom Brocke, J.: Referenzmodellierung – Gestaltung und Verteilung von Konstruktionsprozessen. Logos, Berlin, 2003.
- [FI02] Flick, U.: Qualitative Sozialforschung – Eine Einführung. Rowohlt, Reinbek, 2002.
- [FL05] Froschauer, U.; Lueger, M.: ExpertInnengespräche in der interpretativen Organisationsforschung. In (Bogner, A.; Littig, B.; Menz, W. Hrsg.): Das Experteninterview – Theorie, Methode, Anwendung. VS Verlag für Sozialwissenschaften, Wiesbaden, 2005; S. 223-240.
- [FMT07] Foshay, N.; Mukherjee, A.; Taylor, A.: Does Data Warehouse End-User Metadata Add Value? Communications of the ACM, 50. Jg. (2007), Nr. 11; S. 70-77.
- [GI97] Gleason, D.: Metadata. In (Bischoff, J.; Alexander, T. Hrsg.): Data Warehouse – Practical Advice from the Experts. Prentice Hall, Upper Saddle River, 1997; S. 135-150.
- [Ha04] Haynes, D.: Metadata for Information Management and Retrieval. Facet Publishing, London, 2004.

- [HBO94] Harmsen, A.F.; Brinkkemper, S.; Oei, H.: Situational Method Engineering for Information System Project Approaches. In (Verrijn-Stuart, A.A.; Olle, T.W. Hrsg.): *Methods and Associated Tools for the Information Systems Life Cycle*. North-Holland, Amsterdam, 1994; S. 169-194.
- [He97] Hellmuth, T.W.: *Terminologiemangement – Aspekte einer effizienten Kommunikation in der computergestützten Informationsverarbeitung*. Dissertation, Universität Konstanz, Konstanz, 1997.
- [Kl07] Klesse, M.: *Leistungsverrechnung im Data Warehousing – Entwicklung einer Methode*. Dissertation, Universität St. Gallen, St. Gallen, 2007.
- [Ma00] Marco, D.: *Building and Managing the Meta Data Repository – A Full Lifecycle Guide*. John Wiley & Sons, New York, 2000.
- [Me06] Melchert, F.: *Methode zur Konzeption von Metadatenmanagementsystemen für das Data Warehousing*. Dissertation, Universität St. Gallen, St. Gallen, 2006.
- [MN05] Meuser, M.; Nagel, U.: ExpertInneninterviews: vielfach erprobt, wenig bedacht – Ein Beitrag zur qualitativen Methodendiskussion. In (Bogner, A.; Littig, B.; Menz, W. Hrsg.): *Das Experteninterview – Theorie, Methode, Anwendung*. VS Verlag für Sozialwissenschaften, Wiesbaden, 2005; S. 71-93.
- [RSS00] Rowohl, F.; Schwarz, S.; Strauch, B.: *Entwicklung einer integrierten Metadatenmanagement-Lösung für das Data Warehousing*. Arbeitsbericht BE HSG/CC DWS/04, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2000.
- [Sc00] Schwarz, S.: *Integriertes Metadatenmanagement – Ein Überblick*. In (Winter, R.; Jung, R. Hrsg.): *Data Warehousing Strategie – Erfahrungen, Methoden, Visionen*. Springer, Berlin, 2000; S. 101-116.
- [To99] Tozer, G.: *Metadata Management for Information Control and Business Success*. Artech House, Norwood, 1999.
- [Wi08] Winter, R.; Schmaltz, M.; Dinter, B.; Bucher, T.: *Das St. Galler Konzept der Informationslogistik*. In (Dinter, B.; Winter, R. Hrsg.): *Integrierte Informationslogistik*. Springer, Berlin, 2008; S. 1-16.
- [Wo06] Wortmann, F.: *Entwicklung einer Methode für die unternehmensweite Autorisierung*. Dissertation, Universität St. Gallen, St. Gallen, 2006.

Proaktives Management von Konsistenzbedingungen im Analytischen Performance Management

Stefan Brüggemann

OFFIS e.V.
Escherweg 2
D-26121 Oldenburg
brueggemann@offis.de

Abstract: Im Analytischen Performance Management (APM) definieren Fachanwender zur Überwachung von Zielsystemen benötigte Kennzahlen. Da hier die bereitzustellenden Daten definiert werden, können an dieser Stelle zusätzlich Datenqualitätsaspekte berücksichtigt werden. So lassen sich bereits proaktiv Konsistenzbedingungen festlegen, welche dann von Datenlieferanten berücksichtigt werden müssen. Dadurch wird die Entstehung von Inkonsistenzen vermieden. Die modellierten Konsistenzbedingungen werden ontologiebasiert verwaltet, damit diese selbst untereinander auf Widerspruchsfreiheit geprüft werden können. In diesem Beitrag wird das Vorgehensmodell des APM aufgegriffen und um explizite Modellierung von Datenqualitätsmerkmalen erweitert. Dazu wird mit ProCon ein Modell vorgestellt, welches es erlaubt, Konsistenzbedingungen in multidimensionalen Datenmodellen zu definieren. So wird ein wichtiger Beitrag zum unternehmensweiten Datenqualitätsmanagement geleistet, um präventiv Datenqualitätsmängel bereits bei der Entstehung zu vermeiden.

1 Einleitung

Organisationen setzen zunehmend auf organisationsweites Performance Management zur Kontrolle, Steuerung und Verbesserung der Organisationsleistung. Organisationen streben an, Strategie und Zielsystem zu operationalisieren, um die Organisationsleistung mittels Indikatoren zu überwachen. Mit dem Analytischen Performance Management (APM) wurde eine Methode zur zielgerichteten, kennzahlenbasierten Leistungsmessung in einer Organisation eingeführt.

Das APM stellt besondere Ansprüche an eine Informationslogistik. Informationslogistik befasst sich mit der Informationsbereitstellung mit Entscheidungsbezug. Ziel der Informationslogistik ist es, relevante Informationen in geeigneter Qualität zur Befriedigung der Informationsbedarfe bereitzustellen [WSDB08]. Getreu dem Prinzip des „Garbage in, Garbage out“ [BM99] können inkonsistente, fehlerhafte Daten zu Fehlinformationen und damit zu Fehlentscheidungen führen.

Unter dem Begriff des Datenqualitätsmanagements (DQM) werden die Modellierung, Verarbeitung, Speicherung und Darstellung von Daten mit dem Ziel der Sicherstellung einer hohen Datenqualität zusammengefasst [OWS+08]. DQM kann an verschiedenen Stellen ansetzen, z.B. als nachgelagerte Qualitätssicherung im ETL-Prozess beim Betrieb von Data Warehouse-Systemen. Präventives Datenqualitätsmanagement, beispielsweise durch die Kontrolle von Benutzereingaben, gewinnt immer stärker an Bedeutung. Gerade im aufkommenden Trend der Real Time Enterprises [29] ist eine frühe Datenqualitätskontrolle im Prozess der permanenten Datenintegration notwendig [MN07]. Heutzutage steht die Frage nach einem unternehmensweiten Datenqualitätsmanagement im Vordergrund, wo bereits bei der Entstehung der Daten Fehler vermieden werden können.

In diesem Beitrag wird mit ProCon (*Pro*aktives Management von *Konsistenz*bedingungen) ein neues Konzept vorgestellt, welches es erlaubt, proaktiv Datenqualität im APM zu modellieren. Dieses Konzept ist ein Beitrag zum präventiven DQM, da es die Entstehung inkonsistenter Daten vermeidet. Bereits zum Zeitpunkt der Modellierung des Informationsbedarfs werden Konsistenzbedingungen modelliert. Hierzu wird eine Erweiterung des Vorgehensmodells des APM vorgeschlagen, so dass Fachexperten zur Modellierung der Datenqualität herangezogen werden können.

In Abschnitt 2 wird zunächst der Begriff „Analytisches Performance Management“ (APM) präzisiert, um das Problemumfeld zu konkretisieren. Zur Verdeutlichung wird ein Beispielszenario aus dem Gesundheitswesen gezeigt. Dann werden in Abschnitt 3 grundlegende Begriffe des Datenqualitätsmanagements erläutert und ausgewählte Datenqualitätsmerkmale erläutert. Die multidimensionale Analyse- und Managementsicht analytischer Informationssysteme wird in Abschnitt 4 eingeführt. Dort wird das multidimensionale Datenmodell MADEIRA erläutert, welches Grundlage für das im darauffolgenden Abschnitt eingeführte Metamodell ProCon ist. ProCon erlaubt es, Konsistenzbedingungen explizit in multidimensionalen Datenmodellen zu formulieren. Weiter wird eine Erweiterung des im APM definierten Vorgehensmodells um explizite Berücksichtigung von Konsistenzregeln vorgestellt. Zur prototypischen Umsetzung wird ein modellgetriebener Ansatz vorgestellt, der es erlaubt, aus diesem Metamodell eine domänenspezifische Sprache zu erzeugen. Für diese Sprache lässt sich dann ein Editor generieren, mit dem Konsistenzbedingungen erstellt werden können. Abschnitt 7 fasst den Beitrag zusammen und schließt mit einem Ausblick auf zukünftige Arbeiten ab.

2 Analytisches Performance Management

In [KM06] wurde der Begriff „Analytisches Performance Management“ eingeführt, vom weiter gefassten Begriff des „Performance Management“ abgegrenzt und wie folgt definiert: APM dient der Überwachung und Steuerung einer Organisation durch

- die kontinuierlich wiederkehrende Modellierung der Zielsysteme von Organisation und Organisationseinheiten,
- deren Verknüpfung untereinander und mit den zur Messung der Zielerreichung herangezogenen Indikatoren sowie

- der Überwachung von Indikatorausprägungen und Zielerreichungsgraden in einem analytischen Informationssystem.

Das APM umfasst also nur die Tätigkeiten, die für die Vorbereitung und Durchführung der zielgerichteten, kennzahlen-basierten Leistungsmessung in einer Organisation relevant sind. Ausgeklammert werden zwei weitere wichtige Merkmale des Performance Managements: Die enge Kopplung zwischen Organisationsstrategie und den für die Zielerreichung implementierten Geschäftsprozessen (Prozessorientierung) sowie die fachliche Ausgestaltung des Performance Managements bezüglich präferierter Strategien, Auswahl der Ziele und Indikatoren, betrachteter Geschäftsprozesse sowie verwendeter Managementinstrumente (Managementmethode).

2.1 Vorgehensmodell des APM

Das APM identifiziert ausgehend von der Organisationsstruktur strategische und operationale Ziele, die durch Kennzahlen operationalisiert werden. Dabei wird für vorab definierte Ziele einer Organisation der Zielerreichungsgrad mittels Kennzahlen gemessen und überwacht. Dabei werden folgende Phasen durchlaufen, die in Abbildung 1 dargestellt sind:

- Bei der Modellierung des Zielsystems formulieren Entscheider, ggf. mit Unterstützung durch Fachexperten, ein Zielsystem. Dieses enthält alle relevanten Ziele der Organisation einschließlich ihrer Ursache-Wirkungsbeziehungen. Alle zur Zielerfüllung geplanten Aktivitäten und Geschäftsprozesse werden hier definiert.
- Werden organisationale Rahmenbedingungen geändert, oder wird unzureichende Effektivität der Ziele oder Strategien festgestellt, wird durch das Hinterfragen der Ziele eine Adaption des Zielsystems ausgelöst.
- Im Rahmen der Modellierung des Messsystems identifizieren und spezifizieren Fachexperten, ggf. unterstützt durch Statistiker, die für die Überwachung der Zielerreichung notwendigen Indikatoren und Kennzahlen.
- Durch das Hinterfragen der Indikatoren und Kennzahlen wird die Anpassung der Indikatoren und Kennzahlen ausgelöst, um die Messung der Zielerreichung zu verbessern.
- Im Subprozess Performance Measurement werden im Rahmen der operativen Planung quantifizierbare Zielwerte für die Indikatoren festgelegt. Diese werden für die fortlaufende Überwachung der Zielerreichung genutzt. In der Analysephase werden die Ursachen für die Planabweichungen durch Statistiker und Datenanalysten untersucht.

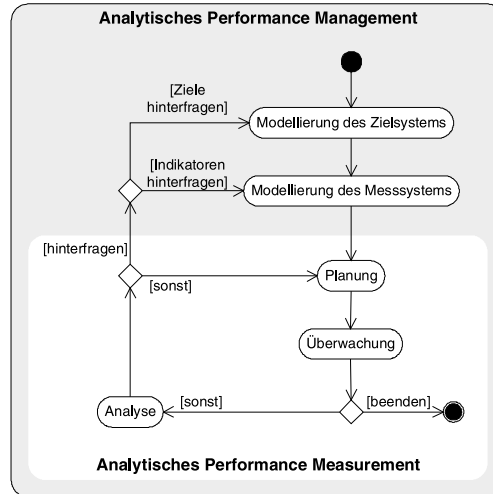


Abbildung 1: Vorgehensmodell des Analytischen Performance Management

Entsprechend der Theorie des organisationalen Lernens wird zwischen den zwei Lernprozessen Anpassungs- und Veränderungslernen unterschieden. Beim Anpassungslernen prüfen Fachexperten, ob sich Indikatoren für die Operationalisierung von Zielen als geeignet herausgestellt haben. Die verwendeten Indikatoren werden hinterfragt und ggf. angepasst, um die Messung der Zielerreichung zu verbessern. Eine mögliche Anpassung ist die Verbesserung der Risikoadjustierung. Bei Änderungen organisationaler Rahmenbedingungen oder unzureichender Effektivität der verfolgten Ziele oder Strategien hingegen wird durch das Hinterfragen der Ziele eine Adaption des Zielsystems ausgelöst. Entscheider und Fachexperten hinterfragen Strategie und Ziele und formulieren diese ggf. neu. Die Effektivität von Zielen im Hinblick auf die Erreichung übergeordneter Ziele und der Organisationsstrategie wird geprüft. Infolge der Neuformulierung werden auch Indikatoren und Kennzahlen durch Fachexperten angepasst.

Werden im Performance Measurement Abweichungen von Indikatoren von ihren Referenzwerten festgestellt, so können diese verschiedene Ursachen haben:

- Verfehlung der Zielerreichung: Entsprechend dem eigentlichen Zweck signalisiert ein Indikator, dass das Ziel, dem der Indikator zugeordnet ist, verfehlt wird.
- Unzureichende Referenzwerte: Der Referenzwert wurde für den Indikator zu restriktiv gewählt.
- Unzureichender Indikator: Der Indikator hat nicht alle relevanten Zusammenhänge repräsentiert und muss entsprechend adjustiert werden, z.B. durch geeignete Einbeziehung von Risikofaktoren in die Berechnungsvorschrift

- Unzureichende Datenqualität: Die Berechnung des Indikators beruht auf Daten unzureichender Qualität.

In diesem Beitrag soll die Ursache der mangelnden Datenqualität besonders betrachtet werden. Insbesondere wird ein Vorgehen eingeführt, wie im Prozessmodell des Analytischen Performance Management Datenqualität explizit berücksichtigt werden kann.

2.2 Beispielszenario: Qualitätsmessung im deutschen Gesundheitswesen

Im Gesundheitswesen wird im Rahmen der externen vergleichenden Qualitätssicherung die Leistung in den deutschen Krankenhäusern erfasst. Die medizinische und pflegerische Qualität wird mittels geeigneter Indikatoren sichtbar gemacht. Dazu dokumentieren alle Krankenhäuser qualitätsrelevante Daten für bestimmte Leistungsbereiche. Diese Daten werden gemäß der Vorgaben des Gesetzgebers (§ 137 Sozialgesetzbuch V) von der BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH [BQS08] analysiert, um beispielsweise die Qualität der Krankenhausversorgung sichtbar und vergleichbar zu machen. Die BQS, gegründet im Jahr 2000 von den Selbstverwaltungspartnern im Gesundheitswesen, leitet und koordiniert die inhaltliche Entwicklung und organisatorische Umsetzung der externen vergleichenden Qualitätssicherung.

Die BQS definiert in Fachgruppen Qualitätsziele und ermittelt für diese Qualitätsindikatoren und Auffälligkeitsgrenzen, mit denen die gesetzten Berechnungsergebnisse bewertet und die Zielerreichung beurteilt werden können. Gemäß dem Vorgehen des Analytischen Performance Managements werden die definierten Qualitätsindikatoren zyklisch hinterfragt und die Auswertungskonzepte fortlaufend überprüft und weiterentwickelt. Um die ca. 2200 deutschen Krankenhäuser miteinander vergleichen zu können, sind folgende Schritte nötig:

- Alle Krankenhäuser erheben qualitätsrelevante Daten und senden sie der BQS
- Die zuvor definierten Kennzahlen werden von der BQS gerechnet und analysiert
- Die Ergebnisse werden den Krankenhäusern berichtet und gegebenenfalls werden Empfehlungen ausgesprochen
- Sollten Auffälligkeitsgrenzen überschritten sein, werden diese Ergebnisse mit den Krankenhäusern in einem „Strukturierten Dialog“ gemeinsam analysiert und Maßnahmen zur Verbesserung umgesetzt.

Dem Aspekt der Datenqualität kommt im Kontext externer vergleichender Qualitätssicherung eine besondere Bedeutung zu. Basieren die durchgeführten Berechnungen und damit die Vergleiche der gemessenen Krankenhäuser auf fehlerhaften Daten, so sind auch die Vergleiche selbst fehlerhaft. Daher ist es einerseits wichtig, bereits bei der Modellierung der Indikatoren und Kennzahlen durch Fachexperten diese auch schon die Modellierung von (In-)Konsistenzen durchführen zu lassen. Andererseits ist es wichtig, nicht nur die Indikatoren und Kennzahlen, sondern auch die Konsistenzbedingungen kontinuierlich zu adjustieren. Dadurch, dass bereits bei der Definition des Informations-

bedarfs die Datenqualität explizit berücksichtigt und modelliert wird, lässt sich die Qualität der an die BQS gelieferten Daten proaktiv steigern.

Da die Krankenhäuser zyklisch neue Anforderungen der BQS erhalten und umsetzen, sind sie als Datenquelle bereits in der Lage, nur qualitativ hochwertige, den Konsistenzbedingungen genügende Daten zu liefern. Gleichzeitig ist die BQS in der Lage, unzureichende Daten von vornherein abzulehnen.

Beispiele für medizinische Dimensionen, die der BQS von Krankenhäusern zur Dokumentation in der Kardiologie geliefert werden, sind ICD-Kodierungen und OPS-Schlüssel. Die Weltgesundheitsorganisation WHO [ICD] definiert mit der International Classification of Diseases (ICD) eine Klassifikation von Erkrankungen. OPS-Schlüssel [Org78] dienen der formalen Bezeichnung von Operationen und Prozeduren im klinischen Kontext.

3 Datenqualität

Private und öffentliche Organisationen beginnen zunehmend, die Qualität der Daten, die sie verarbeiten, als ein wichtiges Wirtschaftsgut zu verstehen. Daher ist eine hohe Datenqualität anzustreben. Es existiert eine Vielzahl von Studien, die verschiedenste Prozentzahlen für fehlerhafte Daten in Organisationsdatenbanken nennen. Viele (Data Warehouse-) Projekte sind gescheitert, und oftmals sind es Datenqualitätsmängel, die zu fehlerhaften Entscheidungen führen. Data Warehouses liegen oftmals als Datenbasis im APM zu Grunde. Es ist von essentieller Wichtigkeit, dass diese Datenbasen genaue, vollständige und konsistente Daten enthalten.

Im Folgenden soll in Abschnitt 3.1 zunächst aufgezeigt werden, wie der Begriff der Datenqualität definiert werden kann und durch welche Merkmale Datenqualitätsmängel beschrieben werden können. In Abschnitt 3.2 werden Konsistenzprobleme explizit dargestellt. Abschnitt 3.3 zeigt, wo Datenfehler entstehen können und Abschnitt 3.4 schließlich zeigt, wie Konsistenzprobleme behoben werden können.

3.1 Datenqualitätsmängel und -merkmale

Die gegenwärtige Literatur verwendet keine einheitliche Definition des Begriffs Datenqualität. Vielmehr wird Datenqualität durch verschiedene Datenqualitätsmerkmale beschrieben. Im Folgenden wird ein Überblick über die gebräuchlichsten Merkmale gegeben. Im Beitrag von Scannapieco et al. [SMB05] werden die Merkmale Konsistenz, Genauigkeit und Vollständigkeit untersucht: Unter dem Begriff der Vollständigkeit werden Nullwertbehandlungen betrachtet, Genauigkeit betrachtet den syntaktischen und semantischen Abstand eines falschen Wertes zu einem richtigen. Konsistenz behandelt die Einhaltung semantischer Regeln, beispielsweise Integritätsbedingungen.

Naumann [Nau02] definiert Genauigkeit einer Datensammlung als den Quotienten der Anzahl korrekter Werte und der Anzahl aller Werte. Müller und Freytag [MF03] definie-

ren Konsistenz Schemakonformität. Sie wird gemessen als Anzahl fehlerfreier Tupel einer Relation in Bezug zur Anzahl aller Tupel einer Relation. Fehlerfreiheit bezieht sich hier auf die Verletzung der syntaktischen Struktur einer Relation. Müller und Freytag definieren die Korrektur domänenspezifischer Datenfehler als größte offene Herausforderung bei der Datenbereinigung. Dies ist darin begründet, dass Wissen bislang nicht ausreichend modelliert wird, um Korrekturwerte für invalide Tupel zu identifizieren. Da sich Ontologien zur Wissensmodellierung eignen wird in dieser Arbeit ein ontologiebasierter Ansatz eingeführt.

Ontologien sind formale explizite Spezifikationen einer gemeinsamen Konzeptualisierung [Gru93]. Sie repräsentieren Konzepte und ihre Beziehungen untereinander. Ontologien enthalten weiter Inferenz- und Integritätsregeln, mit denen Plausibilitätszusammenhänge geprüft werden können.

3.2 Betrachtung von Konsistenzproblemen

Konsistenz bezeichnet die korrekte Kombination der Attributwerte, durch die ein Tupel beschrieben werden kann. Attributwerte können sowohl in linearen Wertebereichen (bspw. Alter) definiert sein, als auch hierarchischen Metadatenkatalogen entstammen. Es wird bei der Prüfung auf Konsistenz hier die Einhaltung semantischer Regeln betrachtet, durch die eine multidimensionale Entität konsistent beschrieben werden kann. In multidimensionalen Datenmodellen können semantische Regeln an Dimensionsmetadaten geknüpft werden. Der Begriff der Konsistenz bezeichnet allgemein die Verletzung semantischer Regeln, die für Daten und deren Attribute definiert sein können. Im relationalen Bezug können dies beispielsweise Integritätsbedingungen sein. Diese Integritätsbedingungen müssen von allen Tupeln einer Datenbank erfüllt sein. Semantische Regeln werden üblicherweise in zwei Kategorien eingeteilt:

- Intra-relationale Konsistenzbedingungen: Diese Bedingungen betreffen einzelne oder mehrere Attribute einer Relation. Diese können sowohl auf Schemaebene als auch auf Instanzebene auftreten. Auf Schemaebene werden Bedingungen definiert, die für alle Tupel dieser Relation gelten, beispielsweise „Sterbedatum jünger als Geburtsdatum“. Auf Instanzebene werden Bedingungen zwischen konkreten Werte-Ausprägungen definiert.
- Inter-relationale Konsistenzbedingungen: Diese Bedingungen definieren Regeln, die Attribute betreffen, die aus verschiedenen Relationen stammen können.

3.3 Auftreten von Datenqualitätsmängeln

In der Vergangenheit wurden vielfältige Maßnahmen entwickelt, die aufgetretene Datenfehler aufspüren und bereinigen können. Datenbereinigung wurde als eines der größten Probleme im Data Warehousing betrachtet [RD00]. Datenprobleme werden klassifiziert nach dem Ort ihres Auftretens:

- Single Source-Probleme auf Schemaebene: Hier wird von einer Datenbasis ausgegangen, in der es unzureichende Integritätsbedingungen und mangelhaftes Schemadesign geben kann. Besondere Probleme sind die Nichteindeutigkeit von Schlüsseln, mangelhafte referentielle Integrität oder Werte außerhalb des zulässigen Wertebereichs.
- Single Source-Probleme auf Instanzebene: Auf Tupelebene innerhalb einer Datenbasis kann es Probleme bezüglich konkreter Ausprägungen der Attributwerte geben. Diese können ungültige, fehlende, widersprüchliche oder deplatzierte Werte sein, aber auch ungültige Attributwertkombinationen. Solche Probleme lassen sich in der Regel nur mit Hilfe domänenspezifischen Wissens identifizieren und beheben.
- Multi Source-Probleme auf Schemaebene: Bei mehreren, sich inhaltlich überlappenden Problemen kann es auf Schemaebene zu Namenskonflikten und strukturellen Konflikten kommen. Ursächlich hierfür sind heterogene Datenmodelle und Schemadesigns.
- Multi Source-Probleme auf Instanzebene: Auch auf Instanzebene kann es bei der Betrachtung mehrerer Datenquellen zu Problemen kommen. Besonders auffällig sind hier Redundanzen und Inkonsistenzen. Benutzen verschiedene Krankenhäuser beispielsweise unterschiedliche Klassifikationssysteme zur Befundung, lassen sich diese Daten nicht oder nur eingeschränkt verglichen.

In diesem Beitrag werden Single Source-Konsistenzprobleme auf Instanzebene betrachtet.

3.4 Edit/Imputation – Systeme

Die Erkennung und Behebung von Konsistenz-Verletzungen ist ein wohlbekanntes Problem. Im statistischen Bereich, beispielsweise bei der Verarbeitung von Fragebögen, werden Regeln (edits) definiert. In einem Fragebogen kann zum Beispiel die ungültige Aussage ICD=I21.1 und OPS=5.728.0 erfasst worden sein. Eine Regel, um einen solchen Fehler zu erkennen, könnte sein: Wenn ICD=I21.1, dann muss OPS \leq 5.72 sein. Nachdem solch ein Fehler erkannt worden ist, muss dieser fehlerhafte Datensatz korrigiert werden. Dazu werden Daten in fehlerhaften Feldern mit korrekten Werten ersetzt (imputation). Im amerikanischen Census Bureau¹ werden demographische Daten mit Hilfe solcher Regeln qualitätsgesichert. Dieses Vorgehen der Fehlererkennung und Ersetzung fehlerhafter Werte wird in der Literatur als Edit-Imputation-Problem bezeichnet.

Die Fellegi-Holt-Methode [FH76] ist ein bekanntes theoretisches Modell [LKM06, SMB05] dieses Ansatzes. Es hat drei wesentliche Ziele:

¹ Das Census Bureau befasst sich mit der Erhebung von Bevölkerungsdaten (siehe www.census.gov, zuletzt besucht am 12. März 2008).

- Die Daten jedes Datensatzes sollen alle Edits erfüllen, indem nur die wenigsten Felder geändert werden. So bleibt ein Großteil der Originaldaten unverändert.
- Ersetzungsregeln sollen automatisch aus den Edits abgeleitet werden. So wird gewährleistet, dass ersetzte Tupel die Edits erfüllen. Weiter wird so die Wartbarkeit verbessert.
- Ersetzte Werte sollen in Hinblick auf die Häufigkeitsverteilung des betrachteten Attributs gewählt werden. So soll nicht nur der Durchschnittswert des Attributs zur Ersetzung verwendet werden, um die Verteilung nicht zu verfälschen.

Winkler et al. [HSW07, Win04] geben eine Einführung in das Themengebiet. Es existieren diverse Implementierungen dieses Modells [WQ03, WP97, MR01], dennoch besteht massiver Forschungsbedarf insbesondere beim Problem der Fehlerlokalisierung.

Es gibt leider einige Einschränkungen im Bezug auf Edit/Imputation Systeme, so dass diese leider nicht bei der Definition des Informationsbedarfes in einer Data Warehouse Umgebung zum Einsatz kommen können:

- Bisherige Implementierungen sind nicht auf Data Warehouses angewandt, so dass es hier insbesondere keine Anpassung des Modells an multidimensionale Datenstrukturen gibt. Durch die Ausnutzung hierarchischer Strukturen lassen sich invalide Tupel besser korrigieren, da nicht nur statistisch ermittelte Werte, sondern semantisch passende gewählt werden können.
- Edit/Imputation-Systeme sind nicht dazu konzipiert, proaktiv in den Prozess der Informationsbeschaffung und die Definition des Informationsbedarfes einzugreifen. Ihre Anwendung findet sich beispielsweise in ETL-Prozessen, wo sie reaktiv Inkonsistenzen in Daten finden und beheben können.
- Regeln (edits) in Edit/Imputation Systemen sind weniger mächtig als Integritätsbedingungen in Informationssystemen, da sie nicht auf einem Datenmodell wie dem Relationalen basieren [SMB05].
- Bei der Definition von Konsistenzbedingungen können Fehler und Widersprüche auftreten, so dass zwei Regeln sich gegenseitig ausschließen. Edit/Imputation Systeme beachten diese Widersprüche bisher nicht.

Um Widersprüche in Konsistenzbedingungen aufzufinden, wird in dieser Arbeit in Abschnitt 5 ein ontologiebasierter Ansatz eingeführt. Ontologien wurden bereits in [LN07] zum DQM bei der Informationsintegration und in [BA07] zur ontologiebasierten Datenvalidierung und -bereinigung eingeführt.

4 Multidimensionale Analyse- und Managementsicht

Analytische Informationssysteme bieten für die explorative Analyse von Daten typischerweise eine multidimensionale, auswertungsorientierte Sicht auf die in einem Data Warehouse integrierten Daten [CG06]. Dabei wird von einzelnen Individuen, Fällen oder Transaktionen (Mikrodaten) abstrahiert. Stattdessen werden aggregierte, beispielsweise

nach Zeit oder Organisationseinheit klassifizierte Daten (Makrodaten) in Form multidimensionaler Datenräume (Datenwürfel, Cubes) analysiert. In Forschung und Praxis gibt es diverse multidimensionale Datenmodelle, die für die explorative Analyse von Daten konzipiert sind. Multidimensionale Datenmodelle sind aus folgenden Gründen als Analyse- und Managementsicht für die im APM zur Modellierung und Überwachung von Messsystemen benötigten Informationen geeignet:

- Die Konzepte und Begriffe multidimensionaler Datenmodelle spiegeln die Sichtweise von Analysten und Entscheidern in Dimensionen und Klassifikationshierarchien wider.
- Die zur Messung der Zielerreichung herangezogenen Kennzahlen bzw. Indikatoren haben einen verdichtenden Charakter.

Das multidimensionale Datenmodell MADEIRA (Modelling Analyses of Data in Epidemiological InteRActive studies) [Wie00] integriert die Konzepte und Begriffe verschiedener Modelle aus den Bereichen On-Line Analytical Processing (OLAP) sowie Scientific & Statistical Databases. Dabei wird ein konzeptioneller Ansatz verfolgt, der unabhängig von der Implementierung in einer konkreten Datenbank (beispielsweise Star Schema oder Snowflake Schema [Leh03]) ist. Der multidimensionale Datenraum bildet die zentrale Struktur für Daten und Berechnungsergebnisse. Aufgrund des verdichtenden Charakters der zur Messung der Zielerreichung herangezogenen Kennzahlen eignet sich die multidimensionale Analyse- und Managementsicht für die Modellierung von Messsystemen und deren Überwachung im Rahmen des APM.

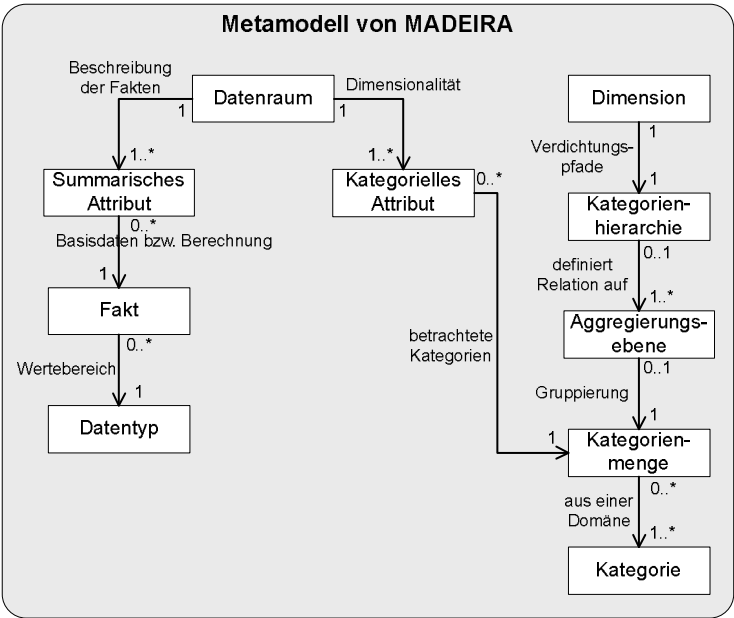


Abbildung 2: Metamodell von MADEIRA

Das Metamodell von MADEIRA ist in Abbildung 2 dargestellt. In MADEIRA wird ein multidimensionaler Datenraum durch summarische und kategorielle Attribute beschrieben. Einzelne Fakten und deren Berechnungen werden durch summarische Attribute definiert. Kategorielle Attribute werden durch mehrere Kategorien gebildet. Kategorien sind elementare Ausprägungen. Nach diesen kann klassifiziert werden (z.B. ICD-Code = 'I21') und sie spannen so den multidimensionalen Datenraum auf. Zugeordnet sind Kategorien zur Aggregierungsebene der Kategorienhierarchie einer Dimension. Dimensionen können beispielsweise 'ICD' oder 'Organisationseinheit' sein. Dimensionen sind unabhängig von betrachteten multidimensionalen Datenräumen und werden organisationsspezifisch modelliert.

5 ProCon: Proaktives Management von Konsistenzbedingungen im APM

Konsistenzbedingungen lassen sich im Analytischen Performance Management direkt an die betrachteten Kennzahlen binden. Da Fachanwender Kennzahlen auf Basis der multidimensionalen Sicht auf das Zielsystem definieren, ist es empfehlenswert, Konsistenzbedingungen direkt in multidimensionalen Datenmodellen zu modellieren.

Rückblickend auf die im Abschnitt 2 beschriebene Phase der Modellierung des Messsystems wird diese hier noch einmal vertieft dargestellt. In dieser Phase werden insbesondere die folgenden Schritte durchgeführt:

- *Datengrundlage prüfen:* Für die ausgewählten Indikatoren wird überprüft, ob die benötigten Daten und Metadaten aus den operativen Applikationen oder aus externen Quellen mit vertretbarem Aufwand beschafft werden können. Es werden also Informationsbedarf und Informationsangebot abgeglichen. Ggf. wird auf Indikatoren verzichtet.
- *Operationalisierung festlegen:* Die Operationalisierung von Zielen durch Indikatoren wird abschließend festgelegt.
- *Berechnungsvorschriften konkretisieren:* Die Berechnung der Kennzahlen, auf denen die Indikatoren basieren, wird exakt in Form von Berechnungsvorschriften beschrieben. Für alle berechneten Kennzahlen wird die komplette Herleitung aus Basiskennzahlen beschrieben.

Werden diese Schritte durchgeführt, so sollte hier auch das Datenqualitätsmanagement berücksichtigt werden. Der Fachexperte kann an dieser Stelle sinnvolle Einschränkungen für die betrachteten Daten und Dimensionen modellieren. Zwischen Kennzahlen, Aggregierungsebenen und Dimensionen lassen sich semantische Regeln definieren, die im Vorgehensmodell des APM berücksichtigt werden müssen. Daher wird folgende Teilkaktivität als Erweiterung der Aktivität der Modellierung des Messsystems definiert:

- *Konsistenzbedingungen modellieren:* Zur Optimierung der Datenqualität lassen sich für Kennzahlen, auf denen Indikatoren basieren, Konsistenzbedingungen definieren.

Es wird nun ein Metamodell eingeführt, mit welchem sich Attributwertkombinationen in multidimensionalen Datenmodellen definieren lassen. Das multidimensionale Datenmodell gibt für das betrachtete Zielsystem die geforderten Daten vor, so dass der Informationsbedarf für das Zielsystem explizit modelliert werden kann. Dimensionen können sowohl aus einzelnen Datenquellen als auch aus Metadatenquellen wie ICD- oder OPS-Kataloge einbezogen werden. Durch das Metamodell können für jedes Zielsystem valide und invalide Verbindungen zwischen Ebenen und Kategorien angegeben werden, so dass bereits bei der Modellierung der ETL-Prozesse Konsistenzbedingungen definiert werden können.

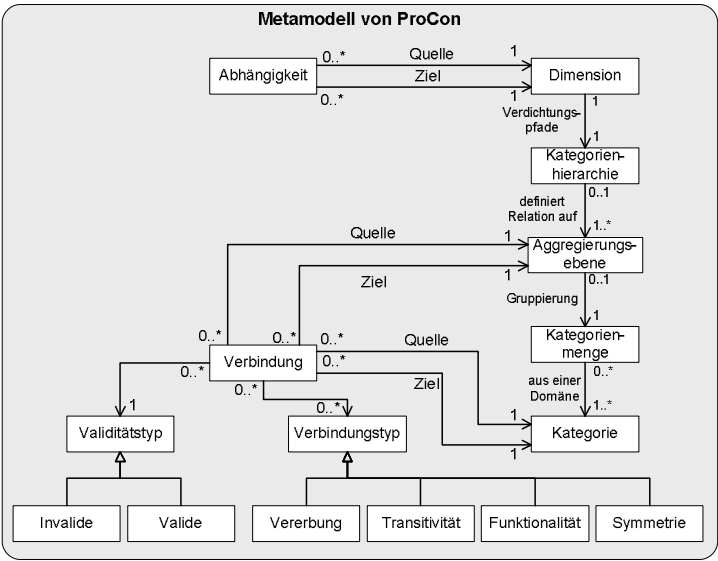


Abbildung 3: Metamodell zur Definition von Konsistenzbedingungen

In Abbildung 3 ist ein Ausschnitt aus dem MADEIRA-Metamodell um Verbindungen erweitert worden. Verbindungen haben immer eine Quelle und ein Ziel und können somit zwei Entitäten verbinden. Kategorien lassen sich mit Kategorien und Ebenen aus anderen Dimensionen verbinden. Gleiches gilt für Ebenen. Für jede Verbindung lässt sich definieren, ob sie einen validen oder invaliden Sachverhalt beschreibt. Valide Sachverhalte beschreiben gültige Attributwertkombinationen und geben somit erlaubte Kombinationen vor. Durch die Modellierung invalider Sachverhalte wird festgelegt, welche Attributwertkombinationen ungültig sind. Solche Daten lassen sich somit direkt bei der Informationsbeschaffung abweisen. Weiter können Verbindungen Eigenschaften zugeordnet werden. Die Eigenschaft „Vererbung“ einer Verbindung zu einer Ebene beschreibt, dass diese Verbindung auch zu allen Kategorien in dieser Ebene gilt. Transitivität gibt an, dass eine Verbindung traversiert werden kann. Eine funktionale Verbindung beschreibt, dass eine Kategorie oder eine Ebene ausschließlich mit der verbundenen gültig ist, und mit keiner anderen. Ist eine solche Verbindung zusätzlich symmetrisch, so ist ausgesagt, dass beide Verbindungsenden ausschließlich mit dem anderen Ende verbunden werden dürfen. Damit Regeln nicht zwischen beliebigen Dimensionen erstellt

werden, sondern nur zwischen solchen, zwischen denen es semantische Zusammenhänge gibt, kann man mit Hilfe des Metamodells solche Abhängigkeiten modellieren. Da Abhängigkeiten jeweils eine Quelle und ein Ziel haben, lassen sich so zwischen Dimensionen semantische Beziehungen modellieren.

Nach der Modellierung der Konsistenzbedingungen werden diese in Ontologiestrukturen überführt. Ontologien lassen sich ideal auf Widersprüche analysieren. Hier können sich gegenseitig widersprechende Konsistenzbedingungen aufgespürt werden. Diese können dann innerhalb der Aktivität *Konsistenzbedingungen modellieren* dem Fachanwender wieder präsentiert werden, so dass diese analysiert und neu adjustiert werden können.

6 Evaluation

Zum Zwecke eines „Proof of Concept“ wurde ein Prototyp entwickelt (siehe Abschnitt 6.1). In Abschnitt 6.2 ist ein Konsistenzszenario exemplarisch dargestellt, welches mit diesem Prototyp modelliert worden ist.

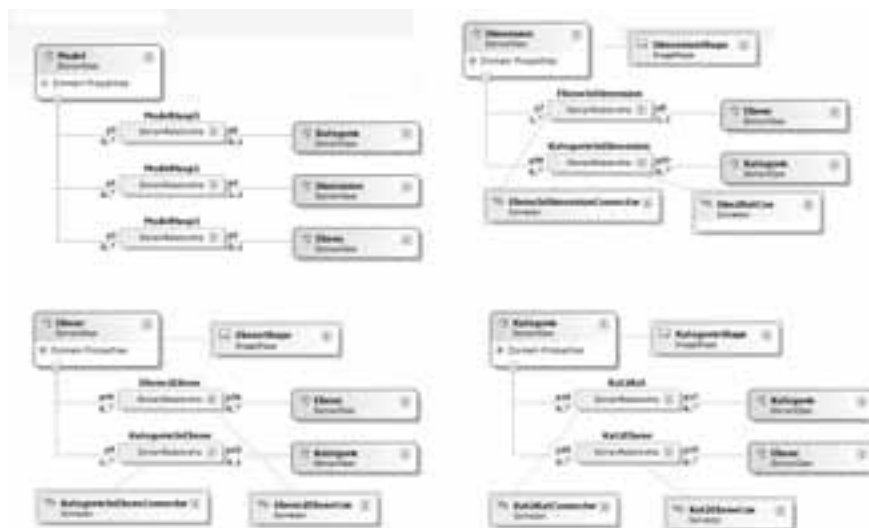


Abbildung 4: Definition einer domänenspezifischen Sprache zur Erstellung von Konsistenzbedingungen

6.1 Prototypische Realisierung

Zur Erstellung der Konsistenzregeln wurde eine domänenspezifische Sprache (DSL) [LKT04] definiert, die in Abbildung 4 dargestellt ist. Auch hier sind wieder die Elemente Dimension, Ebene und Kategorie enthalten. Um multidimensionale Strukturen zu modellieren, lassen sich Kategorien Ebenen und auch Ebenen Dimensionen zuordnen. Ebenso gibt es die Elemente zur Definition von Verbindungen zwischen Kategorien und

Kategorien, Kategorien und Ebenen sowie Ebenen und Ebenen. Diese Sprache wurde mit Hilfe der Microsoft DSL Tools [Cor08] erstellt. Für diese Sprache lässt sich automatisch ein grafischer Editor erzeugen, mit welchem sich dann Konsistenzbedingungen in multidimensionalen Datenmodellen erstellen lassen, die gültige Instanzen der DSL und des Metamodells sind.

Ein Überblick über den modellgetriebenen Ansatz der Modellierung und Erstellung der Konsistenzregeln ist in Abbildung 5 dargestellt. In Teil A ist das Metamodell ProCon dargestellt. Es enthält die relevanten Elemente aus MADEIRA und die Konzepte zur Definition von Konsistenzbedingungen. Dieses Metamodell wird im nächsten Schritt auf eine DSL abgebildet (Teil B). Diese Sprache besteht aus den Entitäten und Beziehungen des Metamodells. Aus dieser DSL lässt sich dann automatisch ein Editor erzeugen (Teil C). Die mit diesem Editor modellierten Strukturen werden dann in eine Ontologie überführt und dort auf Widerspruchsfreiheit geprüft. Auf Grund dieses Vorgehens ist die erzeugte Ontologie sowohl Instanz des Metamodells als auch der domänenspezifischen Sprache.

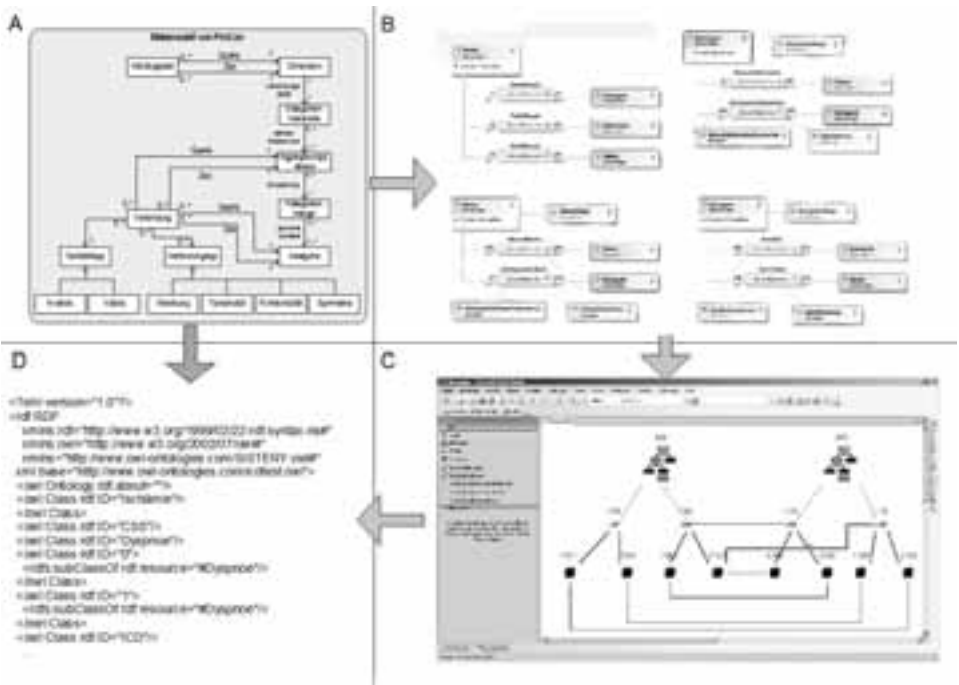


Abbildung 5: Abbildungsschritte: Das Metamodell ProCon wird in eine DSL überführt. Aus dieser wird dann ein Editor generiert. Die hierin erzeugte Struktur wird in eine Ontologie transformiert. Diese Ontologie ist Instanz des Metamodells ProCon.

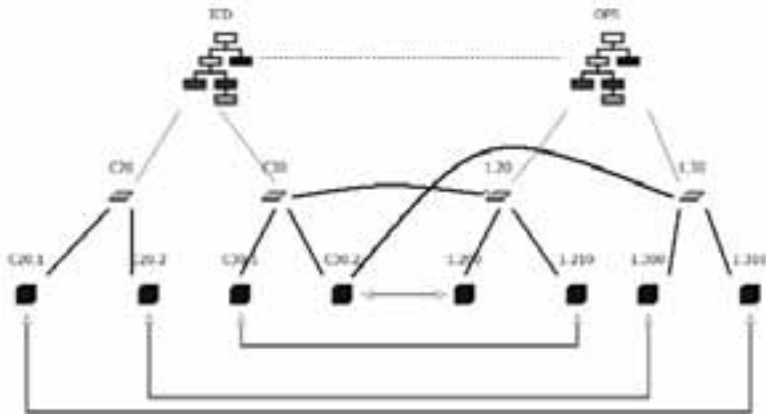


Abbildung 6: Einige mittels des generierten Editors modellierte Konsistenzbedingungen im Kontext der Dokumentation für die Qualitätskontrolle der BQS als exemplarische Anwendung des Ansatzes auf eine konkrete Domäne

6.2 Beispiel

Mit ProCon wurde in dieser Arbeit ein allgemeingültiger Ansatz zur Definition von Konsistenzbedingungen im APM eingeführt. Durch die Verwendung des Metamodells ist der Ansatz auf beliebige Domänen anwendbar. Der realisierte Prototyp wurde genutzt, um die in diesem Beitrag eingeführten Konzepte anhand des Beispielszenarios „Qualitätsmessung im deutschen Gesundheitswesen“ zu evaluieren. Zu diesem Zweck wurden zunächst mit dem auf dem ProCon-Metamodell basierenden Werkzeug die Dimensionen ICD und OPS beispielhaft modelliert.

Abbildung 6 verdeutlicht dieses Szenario. Zwischen ICD und OPS wurde eine semantische Abhängigkeit definiert (gestrichelte Linie). In der Dimension ICD wurden die Ebenen und Kategorien „C20“, „C30“, „C20.1“, „C20.2“, „C30.1“ sowie „C30.2“ modelliert. In der Dimension OPS wurden in der Ebene „1.20“ die Kategorien „1.200“ und „1.210“, in der Ebene „1.30“ die Kategorien „1.300“ und „1.310“ erzeugt. Daraufhin wurden folgende Konsistenzbedingungen definiert: Es ist angegeben, dass der ICD-Code „C30“ nicht in Kombination mit den jeweiligen „1.20“-Werten der anderen Dimensionen auftreten darf (geschwungene Verbindung). Da diese Verbindung auch Grund der hierarchischen Struktur der OPS-Klassifikation und der Nutzung des Verbindungstyps „Vererbung“ zu den Kategorien „1.200“ und „1.210“ vererbt wird, ist der Wert „C30“ auch in Kombination mit „1.200“ und „1.210“ ungültig. Weiter ist modelliert worden, dass die Kombination von „C30.2“ und den Kategorien der OPS-Ebene „1.30“ „1.300“ und „1.310“ ungültig ist (stark geschwungene Verbindung). Im unteren Teil der Abbildung 6 sind einige valide Verbindungen (durchgezogene Linien) angegeben, beispielsweise darf die Kategorie „C20.2“ mit „1.300“ kombiniert werden. Ebenfalls darf „C30.1“ mit „1.210“ in Kombination verwendet werden. Dies ist jedoch eine widersprüchliche Modellierung, da dies im Gegensatz zur vormals definierten Aussage steht, dass „C30“ und alle spezielleren Werte nicht mit den OPS-Werten der Ebene „1.20“

kombiniert werden darf. Dieser Widerspruch lässt sich mit den in Kapitel 3.4 beschriebenen Methoden nicht aufdecken. Dies wird erst durch eine Überführung der ProCon-Modelle in Ontologien ermöglicht. Durch eine Validierung der Ontologie lässt sich dieser Widerspruch erkennen. Damit lässt sich die Definition solcher widersprüchlicher Aussagen vermeiden.

7 Zusammenfassung und Ausblick

In dieser Arbeit wurde mit ProCon ein Metamodell eingeführt, mit welchem Konsistenzbedingungen in multidimensionalen Datenmodellen definiert werden können. Das Vorgehensmodell des Analytischen Performance Management wurde aufgegriffen und um das proaktive Management des Datenqualitätsmerkmals Konsistenz erweitert.

Die in diesem Beitrag eingeführten Konzepte wurden beispielhaft auf das Szenario „Qualitätsmessung im deutschen Gesundheitswesen“ angewandt. Hier wurde gezeigt, daß bekannte Konzepte wie multidimensionale Modellierung in diesem Ansatz genutzt werden. Durch die explizite Verwendung eines Metamodells ist die Anwendung des Ansatzes nicht auf die beschriebene Domäne beschränkt, sondern auf andere Domänen übertragbar.

Durch die konsequente Definition von Konsistenzbedingungen direkt bei der Modellierung multidimensionaler Datenstrukturen durch Fachexperten lässt sich Datenqualität in Organisationen proaktiv verbessern. Wird bei der Modellierung des Informationsbedarfs explizit ausformuliert, welche Daten zulässig und welche unzulässig sind, können Inkonsistenzen von vornherein vermieden werden. Bereits bei der Entstehung der Daten greifen diese Konsistenzbedingungen, so dass Anwender diese Mängel direkt bereinigen können. Wurden Datenqualitätsmängel bislang erst bei der Datenintegration beispielsweise in ETL-Prozessen entdeckt und aufwändig korrigiert, so definiert der hier vorgestellte Ansatz bereits für die Datenentstehung hohe Qualitätsanforderungen.

Der hier vorgestellte Prototyp wird zurzeit erweitert. Besonderes Augenmerk wird auf die Widerspruchsfreiheit der modellierten Bedingungen gelegt. Dazu wird die Struktur in eine ontologiebasierte Repräsentation überführt. Mit Hilfe geeigneter Reasoner (z.B. Racer [HM03]) lässt sich diese Struktur dann validieren.

Im Kontext modellgetriebener Integration kann Datenqualitätsmanagement ebenfalls explizit berücksichtigt werden. Im Sinne eines Top-Down-Ansatzes können Konsistenzbedingungen bereits bei der Modellierung der Integrationservices definiert werden. Diese werden dann bei der Datenintegration direkt genutzt.

8 Literaturverzeichnis

- [BA07] Brüggemann, S.; Aden, T: Ontology Based Data Validation and Cleaning: Restructuring operations for ontology maintenance. In (Hitzler, P.; Sure, Y.; Hrsg.), GI Proceedings 109, Band 1, Jgg. 4 of LNI. GI, 2007.

- [BM99] Bothner, U.; Meissner, F. W.: Data Mining und Data Warehouse: Wissen aus medizinischen Datenbanken nutzen. Deutsches Ärzteblatt, 96, 1999.
- [BQS08] BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH Online. <http://www.bqs-online.de>, 2008. zuletzt besucht am 12.02.2008.
- [CG06] Chamoni, P.; Gluchowski, P., Hrsg.: Analytische Informationssysteme. Springer, 2006.
- [Cor08] Microsoft Corporation: Domain Specific Language Tools. Internet: <http://msdn2.microsoft.com/en-us/vstudio/aa718368.aspx>, 2008. Zuletzt besucht am 12. März 2008.
- [FH76] Fellegi, I. P.; Holt, D: A Systematic Approach to Automatic Edit and Imputation. Journal of the American Statistical Association, 71:17–35, 1976.
- [Gru93] Gruber, T. R: A Translation Approach to Portable Ontologies. Knowledge Acquisition, 5(2):199–220, 1993.
- [HM03] Haarslev, V.; Möller, R: Racer: A Core Inference Engine for the Semantic Web. In 2nd International Workshop on Evaluation of Ontology-based Tools, Seiten 27–36, 2003.
- [HSW07] Herzog, T. N.; Scheuren, F. J.; Winkler, W. E: Data Quality and Record Linkage Techniques. Springer, 2007.
- [ICD] ICD 10: International Statistical Classification of Diseases and Related Health Problems, 10th edition. Geneva: American Psychiatric Association; Jun 1, 1992. World Health Organization.
- [KM06] Koch, S.; Meister, J.: Adaptives Performance Management mit Annotierten Strategy Maps. In (Schelp, J.; Winter, R.; Frank, U.; Rieger, B.; Turowski, K., Hrsg.): Integration, Informationslogistik und Architektur. Proceedings der DW 2006, Jgg. 90, Seiten 13–33, 2006.
- [Leh03] Lehner, W.: Datenbanktechnologie für Data-Warehouse-Systeme : Konzepte und Methoden. Heidelberg : dpunkt, 2003.
- [LKM06] Lenz, H-J.; Koppen, V.; Muller, R.M.: Edits – Data Cleansing at the Data Entry to assert semantic Consistency of metric Data. ssdbm, 0:235–240, 2006.
- [LKT04] Luoma, J., Kelly, S.; Tolvanen, J.-P.: Defining Domain Specific Modeling Languages: Collected Experiences. In Proceedings of the 4th OOPSLA Workshop on Domain-Specific Modeling, 2004.
- [LN07] Leser, U.; Naumann, F.: Informationsintegration. dpunkt.verlag, 2007.
- [MF03] Müller, H.; Freytag, J.-C.: Problems, Methods, and Challenges in Comprehensive Data Cleansing. Bericht, Humboldt Universität Berlin, 2003.
- [MN07] Martin, W.; Nußdorfer, R.: CPM – Corporate Performance Management, Kompendium: Analytische Services in einer SOA, Teil 1: Herstellerunabhängige Beschreibung und Referenzarchitektur. August 2007. [http://www.competence-site.de/soa.nsf/3B1B7B84328E667AC12573960074221A/\\$File/cpm_analytische_services_so_a.pdf](http://www.competence-site.de/soa.nsf/3B1B7B84328E667AC12573960074221A/$File/cpm_analytische_services_so_a.pdf) (zuletzt besucht am 30.11.2007).
- [MR01] Manzari, A.; Reale, A.: Towards a new method of edit and imputation of the Italian Census: a Comparison with the Canadian Nearest-Neighbour Methodology. Presented at the International Statistical Institute Meeting in Seoul, Korea., 2001.

- [Nau02] Naumann, F.: Quality-Driven Query Answering for Integrated Information System. LNCS 2261, 2002.
- [Org78] World Health Organization: International Classification of Procedures in Medicine. 1978.
- [OWS+08] Otto, B.; Wende, K.; Schmidt, A.; Kai Hüner, K.; Vogel, T.: Integrierte Informationslogistik, Kapitel Unternehmensweites Datenmanagement, Seiten 211–230. Springer, 2008.
- [RD00] Rahm, E.; Do, H. H.: Data Cleaning: Problems and Current Approaches. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 23(4):3–13, 2000.
- [SMB05] Scannapieco, M.; Missier, P.; Batini, C.: Data Quality at a Glance. Datenbank Spektrum, 14:6–14, 2005.
- [Wie00] Wietek, F.: Intelligente Analyse multidimensionaler Daten in einer visuellen Programmierungsumgebung und deren Anwendung in der Krebsepidemiologie. Dissertation, Universität Oldenburg, 2000.
- [Win04] Winkler, W. E.: Methods for Evaluating and Creating Data Quality. Information Systems, 29, 2004.
- [WP97] Winkler, W. E.; Petkunas, T.: Statistical Data Editing, Kapitel The DISCRETE Edit System, Seiten 56–62. U.N. Economic Commission for Europe, Geneva, 1997.
- [WQ03] DeWaal, T.; Quere, R.: A fast and simple algorithm for automatic editing of mixed data. J. Official Statist, 19, 2003.
- [WSDB08] Winter, R.; Schmaltz, M.; Dinter, B.; Tobias Bucher. Integrierte Informationslogistik, Kapitel Das St. Galler Konzept der Informationslogistik, Seiten 1–16. Springer, 2008.

Text-Mining-Einsatz zur Bewertung der Reputation von Organisationen und Unternehmen

Paul F. Kirchberg

Studiengang Wirtschaftsinformatik
Berufsakademie Ravensburg
Marienplatz 2
D-88212 Ravensburg
kirchberg@ba-ravensburg.de

Abstract: Auch wenn sich im Web diverse Anbieter finden, die es Benutzern erlauben, Bewertungen über Personen, Produkte, Dienstleistungen sowie auch Organisationen und Unternehmen abzugeben und anderen Benutzern zugänglich zu machen, erscheinen diese Informationen nicht immer repräsentativ. Hinsichtlich einer besseren Einschätzung der Reputation ist daher eine breitere Auswertung von Informationen notwendig, die mehr, aber auch unterschiedliche Arten von Quellen berücksichtigt. Dabei sind Verfahren interessant, die textbasierte Quellen wie Web-Seiten, Blogs aber auch Mails und Forenbeiträge automatisiert hinsichtlich ihrer Reputationsaussage analysieren. Bekannte Verfahren aus dem Web-Reputation-Mining sind jedoch anzupassen, wenn Aussagen über ganze Organisationen und Unternehmen analysiert werden sollen.

1 Motivation

Die Reputation von Personen und Unternehmen spielt bei Geschäftsentscheidungen eine sehr wichtige Rolle. Auch haben Studien gezeigt, dass der Wert eines Unternehmens unmittelbar durch die Unternehmensreputation beeinflusst wird [Gu05]. Dies lässt sich beispielsweise auch direkt im privaten Bereich beim Einkauf über eBay feststellen, auch hier nutzt man gerne Informationen hinsichtlich der Bewertung von Verkäufern für seine eigenen Kaufentscheidungen. Solche Bewertungen finden sich an vielen Stellen im Web. So gibt beispielsweise www.meinprof.de Bewertungen zu Vorlesungen und Professoren oder www.holidaycheck.de sammelt Bewertungen zu Ferienreisen. Die Liste solcher Seiten lässt sich beliebig verlängern. Die Bedeutung einer hohen Einstufung ist dabei so hoch, dass hier natürlich auch bewusst Falscheintragungen von Personen vorgenommen werden, damit diese hinsichtlich ihrer Reputation höher eingestuft werden. Selbst Fernsehsender haben diese Problematik aufgegriffen und zeigen (gerne mit versteckter Kamera), wie sich Hoteliers in den Bewertungslisten selbst sehr gut bewerten. Mit ein wenig Fleiß kann hier jeder eine positive Gesamtbewertung erlangen.

Solche Bewertungsseiten haben jedoch ein zweites, nicht weniger großes Problem. Zufriedene Kunden werden eher selten aktiv, um ihre Zufriedenheit mitzuteilen. Ein unzufriedener Kunde ist enttäuscht, will seine Verärgerung loswerden und teilt seine negativen Erfahrungen tendenziell eher mit, wobei hier teilweise ein Rachegedanke näher liegt als die Freundlichkeit der Informationsweitergabe an andere Benutzer. Für eine objektive Einstufung sind auch Erfahrungen zufriedener Kunden aufzuspüren und in Bewertungen zu integrieren. Zufriedene Kunden reagieren dabei viel mehr auf gezielte Fragen, beispielsweise in Foren und Newsgroups, und geben dabei ihre Erfahrungen weiter.

Dies hat zur Folge, dass bei der Ermittlung einer Reputation mehr Datenquellen zu berücksichtigen sind und vor allem auch, dass nicht nur Bewertungen in einem Punktesystem, sondern Fließtext hinsichtlich einer inhaltlichen Bewertung zu berücksichtigen sind. Damit sind Texte aus verschiedenen Quellen für eine umfassende Bewertung der Reputation in einen Analyseprozess zu integrieren, der in der Lage ist, automatisiert Bewertungselemente hinsichtlich des zu untersuchenden Themas zu entdecken, zu sammeln und zu verdichten. Ein weiterer Grund für die Analyse weiterer Datenquellen wie Newsgroups liegt in der Tatsache, dass die Dinge, die Leute im Internet zueinander sagen, viel delikater sind, als das, was bei einem Interview herauskommt.

Mittels Text-Mining-Verfahren und Web-Reputation-Intelligence wurden hinsichtlich vergleichbarer Fragestellungen im Umfeld von Produktbewertungen schon aussagekräftige Ergebnisse erzielt [NH04]. Probleme bereiten jedoch Situationen, bei denen der hinsichtlich der Reputation zu analysierende Begriff in den betrachteten Texten nur ‚nebenbei‘ erwähnt wurde und der eigentliche Inhalt einschließlich Bewertungselementen andere Bereiche betrifft. Da genau dies bei Reputationsuntersuchungen von Unternehmen und Organisationen sehr häufig der Fall ist, beispielsweise bei reinen Adressangaben oder Kontaktinformationen, sind hier eine Reihe von besonderen Aspekten zu berücksichtigen.

Die Zahl der Anbieter von Reputation-Monitoring-Lösungen ist vor allem im Jahr 2005 drastisch angestiegen. Eher im traditionellen Reputation-Management anzusiedelnde Anbieter wie Factiva, ein Joint Venture von Reuters und Dow Jones, stützen sich vornehmlich auf News-orientierte Medienkanäle mit hoher eigener Reputation wie beispielsweise Financial Times, Zeit oder Spiegel [Gu05]. Eher technisch geprägte Unternehmen mit geringerem finanziellem Rückenwind wie IntelliSeek setzen hingegen auf Consumer-generated Media wie Blogs, Newsgroups und Foren. Hier werden 9000 klassische Medien, 4 Millionen Blogs, 11000 Web-Seiten und 45000 Newsgroups überwacht. RSS-Feeds sind hierbei für den Datenzugriff besonders gut geeignet.

2 Der Reputationsbegriff

Die Reputation eines Unternehmens beschreibt die globale Einschätzung des Unternehmens [Na00]. In diese Einschätzung fließen dabei traditionell gemachte Erfahrungen ein [Ha92], über die stärkere Vernetzung, nicht nur über das Internet, wird jedoch die Verarbeitung von Kommunikationsbotschaften für Unternehmenseinschätzungen immer bedeutender [Ma02]. Dadurch konstituiert sich die eigentliche Reputation in der Gesamtheit aller Personen, die ein Unternehmen auch nur dem Namen nach kennen [GW01].¹

Eine Erhöhung der Reputation kann für ein Unternehmen eine Reihe von Vorteilen, wie beispielsweise eine höhere Kundenbindung [PFT95], geringere Kapitalkosten [BR86], geringere Personalfuktuation [Ca92] und auch höhere Wiederverkaufszahlen und Produktpreise mit sich bringen. Gerade beim Kauf von Produkten, bei denen Qualitätsaspekte nicht dominieren, achten Konsumenten auf Signale, die eine gewisse Unterstützung und Absicherung ihrer Entscheidung bewirken.

Der Ruf eines Unternehmens kann damit als ein immaterieller Vermögensgegenstand interpretiert werden, der einen strategischen und schwer imitierbaren Wettbewerbsvorteil darstellt [Sh03]. Folglich ist eine Erhöhung der Reputation eine wichtige strategische Maßnahme. Hierfür stehen verschiedene Ansätze zur Verfügung wie beispielsweise die Übernahme gesellschaftlicher Verantwortung. Doch wie lässt sich der Erfolg solcher Maßnahmen messen? Die bekanntesten Messinstrumente sind *Fortune's Most Admired Companies* (AMAC/GMAC), *Manager Magazin Imageprofile* und *Reputation Quotient* (RQ), die sich jedoch auf einzelne Stakeholdergruppen fokussieren [ES04]. Da aufgrund der Vernetzung von Unternehmen und Personen neue Kommunikationsformen und Formate entstanden sind, bietet sich daher auch die Entwicklung neuer Messverfahren an, die direkt einzelne Kommunikationsformen hinsichtlich Reputationsbewertungselementen analysieren.

3 Verfahren im Reputation Monitoring

Das erklärte Ziel des Reputation Monitoring (auch Reputation Intelligence genannt) ist die automatisierte Einsichtnahme und das (maschinelle) Verstehen der Meinung von Konsumenten. Software analysiert hier Informationen aus unterschiedlichen Quellen, wobei bevorzugt Web-Quellen wie Foren, Online-Zeitungen, Newsgroups, Blogs jedoch auch Email-Kommunikationen, sofern verfügbar, verwendet werden. Die so zu analysierenden Texte werden dabei mittels Text-Mining-Verfahren bearbeitet, die dabei insbesondere versuchen, Bewertungselemente hinsichtlich eines zu betrachteten Objekts zu finden. Die traditionellen Einsatzgebiete betreffen hierbei Analysen hinsichtlich einzelner Marken oder Produkte.

¹ Mit dem Begriff Reputation verwandt ist der Begriff Image, der jedoch eher als kurzfristig änderbar und mit Kommunikationsbotschaften vom Unternehmen anpassbar einzustufen ist. Reputation ist mehr erfahrungsbasiert und die Kommunikation baut auf Erfahrungen auf [ES04].

3.1 Zentrale Vorgehensweisen

Die Analyse von Web-Quellen wie Foren, Newsgroups und Blogs hinsichtlich der Messung der Reputation von einem zu betrachtendem Objekt lässt sich in einzelne Teilbereiche zerlegen:

- *Citation Count*: Zählen der Häufigkeit einer Objektbezeichnung
Dies wird dabei meist im zeitlichen Verlauf und im direkten Vergleich mit Wettbewerbern betrachtet. Insbesondere kann hierdurch die Auswirkung einer Maßnahme besser verfolgt werden. Auch ist es empfehlenswert, die Erwähnungen unterschiedlich zu gewichten, da Erwähnungen in einer Online-Zeitschrift bedeutender sind als Nennungen in einem Blog, der nur vom Autor selbst gelesen wird.
- *Sentiment Detection*: Messung der Stimmung
Hier wird mittels einer inhaltlichen Analyse die Einstufung der Äußerung vorgenommen um festzustellen, ob hier das Objekt positiv, neutral oder negativ dargestellt wird.
- *Clustering-Verfahren*: Bildung von Begriffsgruppen
Diese ermöglichen, die häufigsten im Kontext eines Objektnamens fallenden Termini zu bestimmen um Beziehungen zu anderen Bereichen und Unternehmen/Produkten und damit auch um Reputationseinstufungen durchführen zu können.
- *Quote-Mining*: Sammeln von Kernaussagen
Ähnlich zu Clustering-Verfahren werden die prägnantesten Zitate aus den Informationsquellen extrahiert, da eine genauere Betrachtung hier lohnenswert erscheint.
- *Quellenklassifikation*
Taxonomien, ein Instrument zur Ordnung von Konzepten ähnlich einer Klassenhierarchie oder einem biologischen Klassifikationsschema, werden eingesetzt, um Nachrichten näher zu kategorisieren, beispielsweise um die Datenquellen in Sparten oder Kanäle wie technische oder Kundenquellen aufzuteilen.

3.2 Architektur

Der Grundaufbau einer Reputation-Monitoring-Lösung entspricht meist dem in Abbildung 2 dargestellten Aufbau. Im Backend arbeiten primär Crawler und andere Datensammler wie RSS-Feed-Aggregatoren auf diversen Dokumententypen. Die so ermittelten Daten werden mittels Mining-Werkzeuge indiziert, so dass die zu einem Suchbegriff gehörenden Dokumente schnell ermittelt werden können.

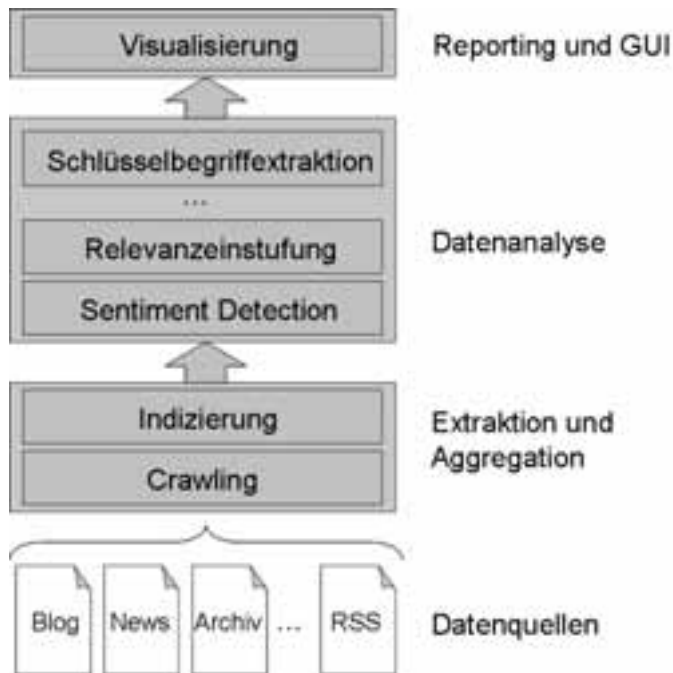


Abbildung 1: Grundaufbau von Reputation-Monitoring-Lösungen

Die zentralen Schritte im Rahmen der Analyse der so ermittelten Daten umfassen POS-Tagging (Part-of-Speech), grammatikalische Analysen, Ermittlung der Einstufung und ihrer Relevanz sowie die Extraktion von Schlüsselphrasen, also besonders häufig vorkommenden Sequenzen von Begriffen. Die Visualisierung umfasst bei einfachen Lösungen eine reine Reputationskennzahl. Hier sind jedoch auch komplexere Zeitreihen- und Datenquellenspartenanalysen möglich, die entsprechend grafisch aufbereitet werden.

3.3 Verfahren

Die Grundlagen der meisten Verfahren stammen aus dem Text-Mining, hier haben sich für textbasierte Analysen Ansätze wie Keyphrase-Extraction, Quote Mining oder Clustering-Verfahren in verschiedenen Einsatzgebieten bewährt. So lassen sich beispielsweise Nachrichten voll- oder semiautomatisch gemäß ihrem Inhalt mit Clustering-Verfahren gruppieren. Auch Verfahren wie das maschinelle Lernen [PLV02], Naive-Bayes-Klassifikationen, Support-Vector-Machines [DLP03] sowie auch neuronale Netze lassen sich hier integrieren.

Gewichtung der Datenquellen

Eine Datenquelle hat mehr Einfluss auf die Gesamtreputation, wenn sie von vielen Personen wahrgenommen wird. Folglich ist gerade bei Web-basierten Quellen die Größe der Leserschaft zu berücksichtigen. Web-Logs bieten hier einige interessante Ansätze an, da typischerweise jeder Blog eine Liste all jener Blogs, die der Autor für interessant und lesenswert befindet, die so genannte Blogroll, enthält. Je häufiger nun ein Blog von anderen empfohlen wird, desto relevanter ist ein Blog einzustufen. Man kann den Reputationswert eines Blogs nun dadurch bestimmen, dass dieser als direkt proportional zur Zahl der eingehenden Blogroll-Links angenommen wird.

Solche Ansätze fanden sich auch schon bei ersten Verfahren für das Ranking von Web-Seiten bei Suchmaschinen, die zum Ranking einer Seite die Anzahl von Links auf eine Seite in die Bewertung integrieren. Aber auch dort hat man schon das Problem erkannt, dass Anbieter mit Zusatzseiten, die nur Verweise auf die eigentliche Seite beinhalten, ihre eigenen Rankings erhöhen [Pa99]. Damit kann auch ein Blog-Schreiber seine eigene Bedeutung erhöhen, indem er eine Reihe von Dummy-Blogs führt, die das primäre Ziel haben, den Hauptblog in der Blogroll zu referenzieren. Hier stellt sich natürlich die Frage, wie häufig ein Blog-Schreiber einen solchen Aufwand überhaupt betreibt. Zur Lösung des Problems können bei Bedarf genau die Verfahren eingesetzt werden, die dieses Problem bei Suchmaschinen angehen. So ist analog zum Ranking bei Google der Page-Rank eines Web-Logs abhängig von der Summe der PageRanks derjenigen Blogs, die in ihrem Blogroll auf ihn verweisen, zu definieren, nicht jedoch von deren bloßer Anzahl.

Neben Blogs sind auch die weiteren zu verwendenden Quellen zu gewichten, um ein aussagekräftiges Gesamtergebnis zu erlangen. Die Gewichtung hängt dabei von einer Reihe von Faktoren ab. Hier ist zu berücksichtigen, wie viele Personen eine Aussage lesen, also auf wie viele Personen eine Aussage Einfluss hat. Auch ist die Glaubwürdigkeit der Quelle bei der Bestimmung der Gewichte zu berücksichtigen, wobei hier auch zwischen der allgemein bzw. selbst wahrgenommen unterschieden werden kann. Diese Gewichtungsfaktoren können dabei nicht allgemein festgelegt werden, insbesondere wenn auch unternehmenseigene Quellen integriert werden. Hier ist eine für eine konkrete Durchführung einer Reputationsbestimmung manuelle Einstellung angebracht, die jedoch bei einer großen Menge an verschiedenen Quellen auch sehr aufwändig werden kann. Wie beschrieben lässt sich bei Blogs die Gewichtung automatisieren. Auch sonst kann man bei Teilbereichen versuchen, Ideen aus dem Pageranking von Suchmaschinen zu übertragen. Weiterhin lässt sich der Aufwand bei der Festlegung von Gewichtungsfaktoren dadurch reduzieren, dass man Quellen zu Gruppen zusammenfasst, die eine vergleichbare Gewichtung besitzen. Mögliche Gruppen sind beispielsweise ‚Online-Nachrichten‘, ‚Beiträge aus allgemeinen Foren‘ oder ‚Beiträge aus produktspezifischen Foren‘.

Polarisation von Aussagen

Bei der Analyse einer konkreten Datenquelle ist primär festzustellen, ob der Inhalt ein Produkt, eine Marke oder eben auch ein Unternehmen neutral, positiv oder negativ darstellt. Diese automatische Einordnung, das *Sentiment Detection*, setzt prinzipiell das Verständnis natürlicher Sprache voraus. Hierfür allgemeine Ansätze zu finden gestaltet sich aktuell nahezu unmöglich, so dass die meisten Verfahren sich inhaltlich auf einen Themenbereich festlegen. Auch sprachliche Besonderheiten wie Ironie und Sarkasmus sowie Metaphern und bildhafte Vergleiche sind kaum mit vertretbarem Aufwand zu integrieren.

Bei den Verfahren können zwei prinzipielle Ansätze unterschieden werden, der rein statistische Ansatz sowie das Natural Language Processing (NLP). Der statistische Ansatz verzichtet dabei auf die Beachtung der grammatikalischen Struktur der Sätze, alle Worte der Datenquelle werden als ungeordnete Wortmenge betrachtet. Diese Wortmenge wird hinsichtlich der statistischen Häufigkeit und der Verteilung von positiven und negativen Wörtern sowie deren konkretes Vorkommen im zu untersuchenden Text analysiert um so die inhaltliche Stimmung abzuleiten. Man erkennt direkt, dass hierfür eine Grunddatenbank notwendig ist, die Aussagen zur Bestimmung der Polarität beinhaltet und damit welche Wörter als positiv und welche als negativ einzustufen sind. Aber auch hier müssen diese Wörter im Kontext betrachtet werden, da beispielsweise das Adjektiv ‚klein‘ bei PDAs als positiv zu bewerten ist, bei Autos eher als negativ eingestuft wird.

Der grundlegende Ansatz beim Einsatz statistischer Verfahren stammt von Turney [Tu02]. Hier werden alle im gesamten Text vorkommenden Begriffe auf einen Stack gelegt. Darauf aufbauend wird für jeden Begriff auf dem Stack mittels der Polarisationszuordnungsdatenbank dessen Polarität ermittelt, die Einzelpolarisationen werden dann abschließend zu einer Gesamtsumme zusammengefasst. Begriffe können dabei aus mehreren Wörtern, so genannten n -Grammen bestehen, wobei n die Anzahl der Wörter in einem n -Gramm darstellt.

Da nun eine Polarisationsdatenbank mit den Zuordnungen von Polarisationen nie vollständig sein kann und insbesondere Probleme im Umgang mit n -Grammen mit sich bringt, hat Turley einen Polarisationserkennungsmechanismus in sein Verfahren eingebaut. Dieser Mechanismus nutzt Suchmaschinen, über die ermittelt wird, wie häufig der zu untersuchende Begriff oder eben das zu untersuchende n -Gramm in textueller Nähe zum Wort ‚exzellent‘ vorkommt. Ist dies häufig der Fall, kann dem Begriff oder dem n -Gramm eine positive Polarität zugeordnet werden. Eine analoge Anfrage bei Suchmaschinen hinsichtlich der textuellen Nähe zu dem Wort ‚schlecht‘ liefert vergleichbare Informationen hinsichtlich einer negativen Polarisierung. Dem Ansatz liegt die Annahme zu Grunde, dass Begriffe mit negativer Interpretation tendenziell eher mit ebenso negativ gefärbten auftreten, während das Gegenteil bei positiven der Fall ist [Tu02].

Alternative statistische Ansätze bedienen sich Mechanismen aus dem Bereich des maschinellen Lernens. So betrachten die Verfahren von Pang (vgl. [PLV02]) Dokumente als Ganzes und zerlegen hierbei die Menge der verfügbaren Texte in eine Trainings- und eine Testmenge. Die Texte der Trainingsmenge dienen dazu, Klassifikatoren zu schulen.

Klassifikatoren sind dabei Algorithmen, die ein Dokument als Eingabe empfangen und auf Basis von dessen Inhalt eine Zuordnung zu einer vorher definierten Menge von Klassen vornehmen. Bekannt sind Klassifikatoren vor allem im Bereich der Spam-Filterung, hier werden Werbe-Mails erkannt und automatisiert aussortiert. Vergleichbar zur Erkennung von Spam-Mails wird hier nun versucht zu erkennen, ob der Inhalt eines Texts eine positive oder negative Meinung darstellt. Damit ein Klassifikator die Zuordnung tatsächlich vornehmen kann, muss er, analog zu Spam-Klassifikatoren, zunächst angelernet werden. Hierbei muss für jedes Dokument, das sich in der Trainingsmenge befindet, manuell eine Zuordnung durchgeführt werden. Dieses Wissen kann der Klassifikator umsetzen, um ein internes Modell mit Kriterien der Zuordnung zu erlernen.

Einer der bekanntesten Klassifikatoren ist der Naive-Bayes, der mit bedingten Wahrscheinlichkeiten arbeitet. Das Verfahren lernt dabei für jedes Wort, beziehungsweise n -Gramm, wie wahrscheinlich es ist, dieses in der Klasse der negativen Meinungen aufzufinden. So könnte nach der Lernphase eine intern erlernte Regel die Form „minderwertig wird in 10 Prozent aller Dokumente mit negativer Meinung eingesetzt“ haben. Dasselbe erfolgt für die Klasse der Begriffe mit positiver Interpretation.

Nach der Lernphase kann das Erlernte auf unklassifizierte Texte angewendet werden. Der Naive-Bayes-Klassifikator zerlegt das zu klassifizierende Dokument in seine einzelnen Begriffe (n -Gramme) und berechnet für jedes n -Gramm die Wahrscheinlichkeit, dass es in der Klasse der positiven oder negativen Meinungen auftritt. Die Wahrscheinlichkeiten werden multipliziert und es ergibt sich eine globale Wahrscheinlichkeit, für jeder der beiden Klassen genau eine. Die Größere der beiden Wahrscheinlichkeiten bestimmt dann das Ergebnis der Klassifikation.

Hinsichtlich einer Verbesserung der Ergebnisse erweitern NLP-Ansätze die statistischen Ansätze um sprachanalytische Elemente. Eine vollständige Zerlegung und tiefe Analyse des Satzbaus ist jedoch im Bereich von Reputationsanalysen nicht durchführbar, da hier eine Massenanalyse verschiedener Quellen erforderlich ist und eine zu detaillierte analytische Zerlegung dadurch aus Performancegründen zumeist ausscheidet. Stattdessen werden so genannte flache NLP-Analysen eingesetzt, die vorgegebene Satzbaumuster auf Texte anpassen [NH04]. So führt beispielsweise IBMs Sentiment Analyzer eine Datenbank mit 3000 Sentiment-Termen, davon 2500 Adjektive und 500 Nomen sowie 120 Sentiment-Mustern [Yi03].

Im ersten Schritt erzeugt IBMs Sentiment Analyzer eine Liste mit allen Begriffen, die im Rahmen der betrachteten Domäne (beispielsweise konkrete Produkte wie Telefone) eine besondere Rolle spielen. Hierzu werden eine Menge spezifischer Dokumente zum Thema verglichen um herauszufinden, welche Begriffe zum Untersuchungsobjekt oft vorkommen. Zum Thema Hochschule extrahiert das Modell beispielsweise die Schlüsseltermine ‚Vorlesung‘, ‚Studiengang‘ oder auch ‚Professor‘. Darauf aufbauend erfolgt die tatsächliche Analyse der Syntax mittels einer Zerlegung der Sätze in atomare Bestandteile wie Adjektive, Verben und Nomen sowie die Darstellung ihres Bezugs zueinander. Ausgehend von einem zerlegten Satz selektiert das Verfahren nun aus den 120 vorgegebenen Sentiment-Mustern die für den Satz passenden Muster. Handelt es sich bei Subjekt oder Objekt um einen der vorher ermittelten Schlüsseltermine, führt der Analyzer die

Analyse fort und prüft vergleichbar zum statistischen Ansatz anhand einer Polarisationsdatenbank die Polarität der Aussage. Das Ergebnis der Satzanalyse ist ein Tupel aus Topic und zugeordneter Meinung.

4 Besonderheiten und Probleme bei der Messung der Reputation von Organisationen und Unternehmen

Unternehmens- oder Organisationsnamen finden sich in sehr vielen Text- und Datenquellen, zumeist jedoch ohne bewertenden Inhalt. So tauchen Firmennamen beispielsweise in Lebensläufen, Adresslisten oder Anschriftentabellen für den Einkauf auf, ohne dass der Inhalt des Dokuments einen inhaltlichen Bezug zu dem Unternehmen hat. Insbesondere sind dadurch Analysen, die die Anzahl der Erwähnungen bei der Bewertung zu stark berücksichtigen hier im Gegensatz zu Produkt- und Markennamen eher ungeeignet [Yi03]. Trotzdem können mittels einer temporalen Betrachtung der Nennungshäufigkeit auch Reputationsbezüge erkennbar sein, da eine Änderung der Nennungshäufigkeit Ursachen haben kann, die nicht unbedingt vom Unternehmen selbst gesteuert werden.

Aus vergleichbaren Gründen sind daher rein statistische Verfahren für die Messung von Unternehmensreputationen ungeeignet. Gerade das Grundverfahren von Turley bietet keine Möglichkeit der Unterscheidung, ob ein Dokument einen Unternehmens- oder Organisationsnamen nur als Nebeninformation oder als Diskussionsgegenstand enthält.

Auch Klassifikatoren, die in vielen Anwendungsgebieten bessere Ergebnisse als das Verfahren von Turley liefern, zeigen bei dieser Fragestellung auch direkt konzeptionsbedingte Schwächen. Hinzu kommt, dass beim Einsatz von Klassifikatoren manuell eine Trainingsmenge bewertet werden muss, was sich schnell als sehr aufwändig erweisen kann. Im Bereich der Messung von Produkt- und Markenreputationen kann man hier auf Quellen zurückgreifen, bei denen Benutzer eine Bewertung nicht nur textuell sondern selbst mit Punkten abgegeben haben. So bieten Anbieter wie beispielsweise Amazon neben der Möglichkeit eine Rezension zu schreiben direkt die Option an, das Produkt mit einem Punktwert auf einer Skala von 1 bis 5 zu bewerten. Dies erlaubt es eine einfache Trainingsmenge aufzubauen, indem man die Textrezensionen als Trainingsmenge verwendet und die dazu gehörende Punktevorgabe dem Lernalgorithmus als Information zum Lernen mitgibt. Der Algorithmus analysiert den Text und gibt seine Prognose bezüglich der Meinung des Autors ab. Anschließend gleicht er sie mit dem mitgegebenen Rating ab. Je größer die Übereinstimmung dabei ist, desto besser erscheint der Klassifikator.

Ein weiterer Nachteil von Klassifikatoren liegt in der Tatsache, dass sie in ihrer Grundform als Ergebnis einer Zuordnung nur die Grundklassen ‚positiv‘ und ‚negativ‘ kennen, evtl. auch noch die Klasse ‚neutral‘. Hier könnten sich Verfahren als vorteilhafter erweisen, die Informationen zum Grad der Polarität auf einer Punkteskala ermitteln. Hinsichtlich der Bewertung der Polarisierung einzelner Wörter ist zudem das jeweilige Zielgebiet mit zu berücksichtigen. So kann das Wort ‚Horror‘ bei einer Hochschule sicherlich mit

einer negativen Bewertung verwendet werden, bei Filmproduktionsunternehmen kann es hingegen evtl. eine neutrale Bedeutung haben, da ein Filmgenre gemeint sein kann.

Diese Problematik ist schon aus anderen Bereichen bekannt. Die Bewertung von Aussagen über Filme lag beispielsweise bei statistischen Verfahren nur knapp über einer willkürlichen Entscheidung, bei Automobilen konnten dagegen mit einfachen Verfahren schon sehr gute Ergebnisse erzielt werden. Der Grund dafür liegt in der schweren Trennbarkeit von inhaltlichen Aussagen und Aussagen über das Objekt. Sofern inhaltliche Aussagen Bewertungsvokabular verwenden, führt eine Interpretation dieser Wörter als Produktbewertung zu falschen Ergebnissen. Durch die reine Betrachtung des Texts als ungeordnete Menge von Wörtern geht dieser Sachverhalt jedoch unter. Erste Analysen zu Dokumenten mit Unternehmensbezug haben hier auch schon gezeigt, dass gerade in diesem Umfeld derartige Vermischungen häufiger auftreten. Folglich sind Konfigurations- und Einflussmaßnahmen in den Messverfahren erforderlich.

Besonders für das Erkennen von neutralen Nennungen von Unternehmensnamen in Anschriften oder Lebensläufen eignet sich das Natural Language Processing sehr gut, wobei hierfür spezielle Muster und Regeln zu definieren sind, die in der Lage sind, diese Fälle von Nennungen zu erkennen und entsprechend einzustufen. Ein einfaches Indiz könnte hier beispielsweise schon die Entfernung einer Nennung zu einem Fließtext sein, dieses Kriterium ist jedoch maximal für eine grobe Vorselektion geeignet. Die grammatikalische Struktur ist in diesem Zusammenhang ein enorm wichtiger Informationsträger und muss mit entsprechenden Mustern berücksichtigt werden, auch wenn dadurch der Aufwand der Analysen erheblich steigt. Als Grundvorgehensweise für die Analyse einzelner Quellen ergibt sich damit die in Abbildung 2 dargestellte Vorgehensweise.

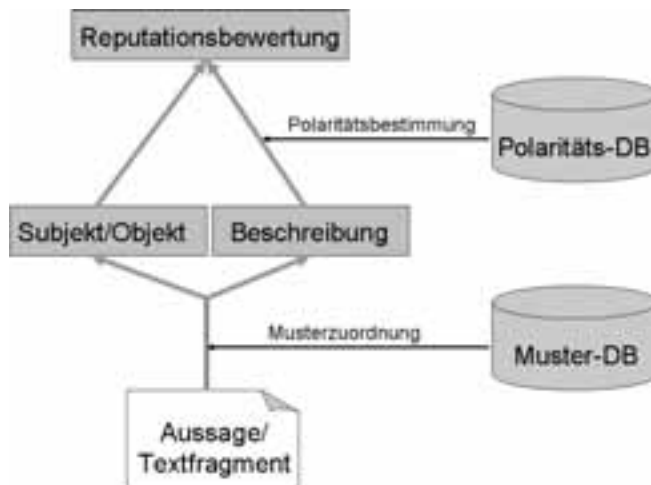


Abbildung 2: Schritte im Rahmen einer Einzelbewertung

Ein weiterer besonderer Aspekt hinsichtlich der Bestimmung der Reputation von Unternehmen und Organisationen liegt in der Tatsache, dass sich die Reputation zum einen hinsichtlich direkter Meinungen über das Unternehmen ergibt, zum anderen Unternehmen auch über ihre Produkte beurteilt werden. Folglich müssen auch Produktmeinungen einbezogen werden, wobei hierfür die Produkte zu nennen sind, zum anderen auch die Bedeutung einzelner Produkte hinsichtlich der Gesamtreputation festzulegen ist. So können negative Aussagen über ein Produkt, welches nur eine geringe Bedeutung für das Unternehmen und in der Öffentlichkeit hat, schwächer gewichtet werden, Aussagen über zentrale, kunden- und öffentlichkeitswirksame Produkte sind dagegen stärker zu berücksichtigen. In diesem Zusammenhang müssen die Verfahren in der Lage sein, Aussagen über Produkte und Aussagen über das Unternehmen an sich zu trennen. Auch hierfür sind nur NLP-Ansätze praktikabel.

5 Die Datenbasis

Nicht jedem, der ein eigenes Reputation Monitoring durchführen möchte, steht eine Datenbasis zur Verfügung, wie sie beispielsweise das auf Reputationsmessungen spezialisierte Unternehmen Factiva besitzt. Eine relativ einfache, wenn auch nicht sehr zielgerichtete Variante ist die Verwendung allgemein zugänglicher Suchmaschinen, die einem Web-Seiten für Suchbegriffe liefern. Der Vorteil dieser Variante ist, dass hierdurch Texte aus allen Bereichen in das Unternehmens-Reputation-Mining einfließen und damit ein breites Spektrum an Quellen integriert wird. Je breiter das Spektrum ist, desto aufwändiger sind jedoch die Muster für die NLP-Verfahren zu definieren. Ein weiterer großer Nachteil beim Einsatz von Suchmaschinen ist, dass nur eine Momentaufnahme abrufbar ist. So werden Forenbeiträge und Online-Artikel, die ein gewisses Alter übersteigen, meist automatisch gelöscht.

Aus diesen Gründen kann es empfehlenswert sein, spezielle Foren, Newsgroups, Blogs aber auch Onlinezeitungen und –nachrichten sowie weitere Informationsanbieter gezielt kontinuierlich abzufragen. Hierdurch können zum einen Regeln definiert werden, die auf das Format der Datenquelle ausgelegt sind, zum anderen können die Daten historisiert in einer Art Meinungs-Data-Warehouse abgelegt werden, wodurch auch zeitbezogene Analysen möglich sind.

Eine weitere wichtige Datenquelle für die Messung der Reputation eines Unternehmens oder einer Organisation sind die Dokumente, die direkt mit Geschäftspartnern ausgetauscht werden. Die hinsichtlich einer Reputationsbeurteilung am einfachsten zu beurteilende Informationsart und dadurch auch schon vielfach eingesetzte Variante ist der Versand von Fragebögen an Partner mit gezielten Fragen hinsichtlich der Beurteilung der Beziehung. Für übergreifende Aspekte ist dieser Ansatz neben anderen Nachteilen nicht geeignet, eine Integration in die Gesamtreputationsbeurteilung erscheint hier natürlich trotzdem sinnvoll. Emails, unternehmenseigene Foren, Kundenportale mit Bemerkungsmöglichkeiten aber auch Telefonprotokolle bieten hier erweiterte Möglichkeiten. Auch diese Informationen können, evtl. zusammen mit den Web-basierten Daten, in ein Auswertungs-Data-Warehouse eingespielt werden, auf dem dann die Verfahren ihre Analy-

sen durchführen. Auch ist der Faktor Ironie bei diesen Quellen von eher geringfügiger Bedeutung.

Ein weiterer Vorteil bei der Berücksichtigung solcher unternehmenseigener Quellen ist die inhaltlich klare Abgrenzung, wodurch sich viel einfacher erkennen lässt, ob der Inhalt hinsichtlich einer Reputationsbewertung Einfluss hat. Insbesondere ist im Gegensatz zu Web-Dokumenten vorgegeben, dass sich die Aussagen um keine anderen Nebenschauplätze drehen. Eine Extraktion der Topics samt deren Bezug untereinander ist somit im Großteil der Fälle im Gegensatz zu Web-Dokumenten oder auch Online-Zeitschriften und -Nachrichten zumeist überflüssig.

Ein großer Vorteil bei der Integration unternehmenseigener Datenbanken mit Texten zu Kundenrückmeldungen liegt weiterhin in der Tatsache, dass diese, evtl. auch direkt bei der Erfassung, mit semantischen Zusatzinformationen versehen werden können, die dann für Auswertungen verschiedenster Art, aber eben auch für das Reputation Mining verwendet werden können. Mechanismen aus dem Semantic Web einschließlich Tagging, hierarchischer Strukturen und Taxonomien vereinfacht die Interpretation des Inhalts für automatisierte Verfahren enorm.

Natürlich hat eine eher breit angelegte Integration von Informationsquellen eine viel größere Aussagekraft. Je mehr Aussagen einbezogen werden, desto repräsentativer sind die Ergebnisse. Hier ist jedoch zusätzlich zu berücksichtigen, dass verschiedene Quellen unterschiedlich zu gewichten sind. So haben Veröffentlichungen in Online-Zeitschriften meist eine größere Leserschaft, wodurch die Auswirkungen von Aussagen gravierender sind. Auch ist die Glaubwürdigkeit höher als bei Blogs, weshalb auch viel gelesene Blogs hinsichtlich der Bewertung nicht einfach gleich eingestuft werden können. Daher ist eine situationsbezogene Gewichtung am geeignetsten, die für einzelne Fragestellungen konfigurierbar sein sollte.

6 Interpretation der Kennzahlen

Die Durchführung einer Reputationsanalyse auf vorhandenen, integrierbaren Datenquellen liefert im einfachsten Fall einen Reputationswert auf einer Skala, die Analyse beschreibt damit eine Messung, das Ergebnis ist eine Maßzahl. Auch wenn die berühmte Aussage „was man nicht messen kann, kann man nicht kontrollieren“ von Tom De Marco [De04] unbestritten ist, ist die Aussagekraft einer solchen Maßzahl eher beschränkt. Zum einen ist zu beachten, dass gerade bei Quellen, die Aussagen einzelner Personen enthalten, eher negative Meinungen als positive Äußerungen erscheinen. Wie schon anfangs erwähnt, neigen unzufriedene Personen mehr dazu, ihre Unzufriedenheit bekannt zu machen, als das zufriedene ihre Situation darzustellen bereit sind. Folglich ist bei jeder Datenquelle auch als Bewertungsfaktor zu berücksichtigen, in wie weit die Quelle aufgrund eigenständiger Aussagen bzw. aufgrund von konkreten Fragestellungen mit Inhalt gefüllt wurden. Bei der zweiten Variante kann hier von einer größeren Neutralität ausgegangen werden.

Das Messen liefert jedoch insgesamt erst dann ein auswertbares Ergebnis, wenn man weiß, was ‚normal‘ ist. Folglich werden Vergleichszahlen benötigt, damit man die Messergebnisse richtig einordnen und einen Abstand erkennen kann, wodurch dann eine Metrik definierbar wird.²

Eine für Unternehmen und Organisationen in Verbindung mit einer vorhandenen Reputationsanalysemethode direkt einsetzbarer Auswertungsmöglichkeit ist eine zeitbezogene Gegenüberstellung von Reputationsmesswerten, die die Betrachtung der Entwicklung der Repräsentation erlaubt. Dadurch können Trendanalysen durchgeführt werden sowie bei negativen Entwicklungen Gegenmaßnahmen veranlasst werden. Auch kann hiermit ermittelt werden, ob eine konkret durchgeführte Maßnahme auch die gewünschte Wirkung in der Öffentlichkeit hat.

Interessanter ist natürlich ein direkter Vergleich zu Mitbewerbern oder anderen Organisationen. Hierfür ist natürlich die Reputationsmaßzahl des Mitbewerbers erforderlich. Da diese entweder nicht öffentlich zur Verfügung steht oder die Ermittlung der durchgeführten Bewertung nicht nachvollziehbar ist, könnten hier eigene Messungen hinsichtlich der Reputation helfen. Werden jedoch bei der Messung der eigenen Reputation interne Dokumente wie Mails und Kundenreports eingesetzt, ist ein direkter Vergleich kaum möglich, da diese Dokumente für andere Unternehmen und Organisationen typischerweise nicht zur Verfügung stehen.

Hinsichtlich einer Vergleichsmessung können folglich nur allgemein zugängliche Informationen verwendet werden. Hierdurch ist dann zum einen ein Unternehmens- und Organisationsvergleich möglich, zum anderen erlaubt dies auch die Feststellung, wie der Unterschied der allgemeinen Reputation zur Reputationsmesszahl mit direkterem Kundenbezug, also bei der auch Mails und Kundenreports in die Messung einbezogen werden, ist. Bestehen hier Unterschiede, kann man im Falle einer negativeren öffentlichen Meinung im Vergleich zu Meinungen mit direktem Kundenbezug versuchen, diese Unterschiede zu analysieren und im Bedarfsfall gezielt Maßnahmen im Rahmen der Öffentlichkeitsarbeit veranlassen. Auch ist evtl. eine Übertragung auf Mitbewerber und andere Organisationen möglich, wodurch die Vergleichbarkeit sich erhöht.

Hinsichtlich eines direkten Vergleichs mit Mitbewerbern und Organisationen ist zudem zu beachten, dass die Anzahl der gefundenen Dokumente in Blogs, Foren, Web-Seiten und Online-Nachrichten in Beziehung zur Unternehmensgröße zu sehen ist, da große Unternehmen allein schon durch die größere Mitarbeiterzahl häufiger genannt werden. Folglich sind im Rahmen eines Citation Count (vgl. Abschnitt 3.1) auch Abwägungen diesbezüglich zu integrieren. Dies gestaltet sich jedoch schwierig, da die Anzahl an Nennungen eines Unternehmens oder einer Organisation von vielen weiteren Faktoren abhängen kann.

² Auch wenn der Begriff Metrik inzwischen sehr frei innerhalb der Bestimmung von Messwerten in unterschiedlichsten Themenstellungen verwendet wird, ist hier der Begriff in seiner ursprünglichen mathematischen Bedeutung als Abstandsmaß zu verstehen.

7 Zusammenfassung und Fazit

In einer 2003 durchgeführten Befragung von über 250 Vorstandsvorsitzenden haben zwei Drittel der Befragten geäußert, dass Reputationsmanagement in ihren Verantwortungsbereich fällt. Nur ein Zehntel gab an, dass dies Aufgabe der PR- oder Kommunikationsabteilung des Unternehmens sei [HQW06]. Um dieser Verantwortung nachzukommen, werden Informationen benötigt, die Aussagen hinsichtlich der aktuellen Situation erlauben sowie Kontrollmechanismen für durchgeführte Maßnahmen integrieren. Neben klassischen Marktforschungsaktivitäten bieten sich aufgrund der Masse an elektronischen Quellen automatische Verfahren, die neben allgemein zugänglichen Daten aus Web-Seiten, Foren oder Blogs auch interne Quellen wie Kundenportale, interne Foren oder Telefonprotokolle berücksichtigen.

Neben den Besonderheiten hinsichtlich der Reputationsinformationsgewinnung aus diesen Quellen ergeben sich für Unternehmen und Organisationen eine Reihe von Besonderheiten, die einer Erweiterung der klassischen Analyseverfahren verlangen. Hier sind sowohl bei der Analyse der Dokumente, bei der Gewichtung von Einzelaussagen und vorhandener Quellen als auch bei der Durchführung von Vergleichen eine Reihe von Aspekten zu beachten, die es zukünftig noch genauer zu analysieren gilt.

8 Literaturverzeichnis

- [BR86] Beatty, R. P.; Ritter, J.R.: Investment Banking, Reputation and Underpricing of Initial Public Offerings. In: Journal of Financial Economics, Vol. 15, Nr. 1/2, 1986, S. 212-232.
- [Ca92] Caminiti, S.: The Payoff from a Good Reputation. In: Fortune, Vol. 125, Nr. 3, 1992, S. 49-53.
- [De04] De Marco, T.: Was man nicht messen kann ... kann man nicht kontrollieren. 2004, Mitp.
- [DLP03] Dave, K.; Lawrence, S.; Pennock, D.: Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. In: Proceedings of the 12th World Wide Web Conference, 2003, S. 519-528.
- [ES04] Eberl, M.; Schwaiger, M.: Die wahrgenommene Übernahme gesellschaftlicher Verantwortung als Determinante unternehmerischer Einstellziele – ein internationaler kausalanalytischer Modellvergleich. Ludwig-Maximilians-Universität München, Schriften zur Empirischen Forschung und Quantitativen Unternehmensplanung, Heft 20, 2004, München.
- [Gu05] Guyer, A.: Eine starke Unternehmensreputation schafft Shareholder Value. 2005, factiva.com/collateral/files/whitepaper_reputationvalue_F-2236_GE.pdf
- [GW01] Gotsi, M.; Wilson, A. M.: Corporate Reputation: Seeking a Definition. In: Corporate Communications, Vol. 6, Nr. 1, 2001, S. 24-30.
- [KF03] Korn/Ferry International: Annual Board of Directors Study. 2003, <http://www.kornferry.com/Library/ViewGallery.asp?CID=581&LanguageID=1&RegionID=23>

- [Ha92] Hall, R.: The Strategic Analysis of Intangible Resources. In: Strategic Management Journal, Vol. 13, Nr. 2, 1992, S. 135-144.
- [HQW06] Heyer, G.; Quasthoff, U.; Wittig, T.: Text Mining: Wissensrohstoff Text. 2006, W3I.
- [Ma02] Mahon, J.F.: Corporate Reputation: A Research Agenda Using Strategy and Stakeholder Literature. In: Business and Society, Vol. 41, Nr. 4, 2002, S. 415-445.
- [Na00] Nakra, P.: Corporate Reputation Management: ‚CRM‘ with a strategic twist. In: Public Relations Quarterly, Vol. 45, Nr. 2, 2000, S. 35-32.
- [NH04] Nigam, K.; Hurst, M.: Towards a Robust Metric of Opinion. In: American Association for Artificial Intelligence Spring Symposium on Exploring Attitude and Affect in Text, 2004.
- [Pa99] Page, L.; Brin, S.; Motwani, R.; Winograd, T.: The Pagerank citation ranking: Bringing order to the Web. <http://dbpubs.stanford.edu/pub/1999-66>.
- [PLV02] Pang, B.; Lee, L.; Vaithyanathan, S.: Thumbs up? Sentiment Classification Using Machine Learning Techniques. In: Proceedings of the EMNLP'02, 2002.
- [PFT95] Preece, S.; Fleisher, C.; Toccacelli, J.: Building a Reputation Along the Value Chain at Levi Strauss. In: Long Range Planning, Vol. 28, Nr. 6, 1995, S. 88-98.
- [Sh03] Shamsie, J.: The Context of Dominance: An Industry-Driven Framework for Exploiting Reputation. In: Strategic Management Journal, Vol. 24, Nr. 3, 2003, S. 199-215.
- [Tu02] Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In: Proceedings of the Meeting of the Association for Computational Linguistics, 2002, S. 417-424.
- [Yi03] Yi, J.; Nasukawa, T.; Bunescu, R. C.; Niblack, W.: Sentiment Analyzer: Extracting Sentiments about a Given Topic Using Natural Language Processing Techniques. In: Proceedings of the IEEE Conference on Data Mining, 2003, S. 427-434.

Ein Konzept zur automatisierten Klassifizierung von Informationen für das Information Lifecycle Management – dargestellt am Beispiel des SAP NetWeaver Business Intelligence

Monique Kosler*, Mauricio Matthesius**, Dirk Stelzer***

*IBM Deutschland
Rathausstraße 7
09111 Chemnitz
Monique.Kosler@de.ibm.com

**sones GmbH
Eugen-Richter-Straße 44
99085 Erfurt
mauricio@sones.de

***Fachgebiet Informations- und Wissensmanagement
TU Ilmenau
Postfach 100565
98684 Ilmenau
Dirk.Stelzer@tu-ilmenau.de

Abstract: Information Lifecycle Management (ILM) hat zum Ziel, Informationen zu klassifizieren und entsprechend ihrer Klasse auf dem jeweils günstigen Speichermedium zur Verfügung zu stellen. Unser Beitrag konzentriert sich auf die zentrale Teilaufgabe des ILM, die Informationsklassifizierung. In der Unternehmenspraxis muss diese zumindest teilweise automatisiert werden. Zunächst ordnen wir die Informationsklassifizierung mit Hilfe eines Vorgehensmodells in das ILM ein. Auf der Grundlage der Vorarbeiten von Chen entwickeln wir ein Konzept zur automatisierten Klassifizierung von Informationen in einem Data Warehouse. Anhand einer prototypischen Implementierung im Rahmen eines SAP-BI-Projektes demonstrieren wir die praktische Anwendung des Konzeptes und erörtern, welche Verbesserungsmöglichkeiten dabei identifiziert wurden.

1 Einleitung

Ein großer Teil der strukturierten Informationen eines Unternehmens befindet sich innerhalb von Datenbanken, beispielsweise in Data Warehouses [Sh06]. Im Data-Warehouse-Bereich ist mit einem starken Informationswachstum zu rechnen, da die Bedeutung von Data Warehouses für Unternehmen stark zunimmt [Sha06], [WW07].

Das Informationswachstum führt zu hohen Speicher- und Administrationskosten [Ab06], [Al01], [TS07]. Die Kosten für die Administration der Datenträger bzw. Datenspeicher, wie beispielsweise die Datensicherung und Spiegelung der Informationen sowie die sichere Aufbewahrung, sind um das vier- bis achtfache höher als die Beschaffungskosten der Datenträger [Gr03] selbst. Zudem müssen Unternehmensinformationen aufgrund der gesetzlichen Anforderungen für längere Zeiträume sicher aufbewahrt werden [Sh06], was zu einer zusätzlichen Erhöhung der Kosten führt.

Information Lifecycle Management (ILM) zielt darauf ab, Informationen¹ zu bewerten, zu klassifizieren und kostengünstig zu verwalten [MHP05], [TS07]. Informationen, die einen höheren Wert für ein Unternehmen aufweisen, müssen von Informationen mit geringerem Wert getrennt [Sh06], [MW99] und den Unternehmensanforderungen entsprechend auf dem jeweils sinnvollsten Speichermedium bereitgestellt und verwaltet werden [Ab05], [Bh05], [MHP05], [Mo06]. Als schwierig erweist sich die Bewertung und anschließende Klassifizierung von Informationen [Ab06], [Bh05], [Ch05], [Me04], [TS07], [YB02]. Hierfür werden zahlreiche Angaben benötigt, wie beispielsweise die Anzahl der Lese- und Schreibzugriffe auf diese Objekte [Ch05], [EI03]. Zusätzlich sind gesetzliche Anforderungen zur Aufbewahrung der Informationen [Sh06] sowie Gewohnheiten der Nutzer [Me04], [Sha06] hinsichtlich der Informationsnutzung von Relevanz. Die Erhebung, kontinuierliche Aktualisierung und Auswertung dieser Bewertungsinformationen ohne automatisierte Unterstützung ist praktisch kaum möglich und wirtschaftlich nicht sinnvoll. Aus diesem Grund muss die Informationsklassifizierung zumindest teilweise automatisiert erfolgen.

In diesem Beitrag entwickeln wir ein Konzept zur automatisierten Klassifizierung von Informationen für das Information Lifecycle Management und stellen dessen Anwendung beispielhaft an SAP NetWeaver Business Intelligence (SAP BI) dar. Für die Bewertung bzw. Klassifizierung von Informationen in Data-Warehouse-Systemen sind unseres Wissens bisher keine vergleichbaren Konzepte publiziert worden.

2 Vorgehensmodell des ILM

Anhand des in Abbildung 1 dargestellten Vorgehensmodells [MS08] werden zunächst wesentliche Funktionen des ILM kurz beschrieben.

¹ Tatsächlich werden im Rahmen des ILM Daten bzw. Dateien analysiert und bewertet. Da in den meisten Publikationen der Betrachtungsgegenstand jedoch als Information bezeichnet wird, verwenden auch wir diesen Begriff.



Abbildung 1: Vorgehensmodell des ILM

In Phase eins, „Analyse der Systemlandschaft“, werden in der Teilphase „Ermittlung der Systeme“ zunächst die in einem Unternehmen vorhandenen Data-Warehouse-Systeme ermittelt und untersucht. Dabei sind insbesondere die Verbindungen² zwischen den Systemen sowie die durch die Systeme verarbeiteten Informationen von Interesse, die in den Teilphasen „Ermittlung der Verbindungen“ und „Ermittlung der verarbeiteten Informationen“ identifiziert werden.

Die physische Größe, die Anwender und das Informationswachstum der Data-Warehouse-Systeme werden in Phase zwei, „Untersuchung relevanter Systeme“, ermittelt. Größe und Wachstum sind wichtige Indikatoren dafür, welche Systeme im Hinblick auf eine Kostenreduktion zuerst untersucht werden sollten. Des Weiteren werden die Anwender der jeweiligen Systeme ermittelt, da diese für die dritte Phase, die „Informationsklassifizierung“, von Bedeutung sind. In der Teilphase „Festlegung von Klassifikationskriterien“ werden für die Informationsklassifizierung relevante Kriterien festgelegt, wie beispielsweise die Zugriffshäufigkeit auf die Informationen sowie das Anwender- und Administratorenwissen über die Systeme und die darin enthaltenen Informationen. Vorgehensweise sowie Methoden zur Messung der relevanten Kriterien werden in der Teilphase „Operationalisierung von Messmethoden und -indikatoren“ bestimmt. Anschließend erfolgt die „Erhebung der Messwerte“ und die „Interpretation und Dokumentation der Ergebnisse“. In diesen beiden Teilphasen werden die Informationen entsprechend der festgelegten Kriterien und Methoden bewertet und klassifiziert.

² Hierbei ist von besonderem Interesse, welche Data-Warehouse-Systeme Informationen aus anderen Data-Warehouse-Systemen aufnehmen, verarbeiten und erneut an andere Data-Warehouse-Systeme weitergeben. So wird deutlich, ob bestimmte Informationen auf unterschiedlichen Systemen redundant gehalten werden und ob eine mehrfache Archivierung dieser Informationen innerhalb der folgenden Phasen des Vorgehensmodells notwendig und sinnvoll ist.

In der Literatur werden unterschiedliche Kriterien zur Bewertung von Informationen herangezogen. Nach unserer Auffassung eignen sich beispielsweise die erzielten monetären Erträge, die aus der Nutzung der Informationen entstehen, oder die entstandenen Kosten, die bei der Sammlung, Akquisition und Erstellung der Informationen anfallen [MW99], eher schlecht. Aufgrund der in der Regel komplexen Verwendung von Informationen lässt sich ein direkter Bezug zu Umsatz- oder Kostengrößen nur sehr schwer herstellen. Wir bevorzugen den Nutzungsgrad als Kriterium für die automatisierte Bewertung von Informationen im Rahmen des ILM. Der Nutzungsgrad gibt an, in welchem Maße eine Information zur Unterstützung von Entscheidungen herangezogen wurde. Er bestimmt sich danach, wie oft und wann der Zugriff auf eine Information erfolgte. Wie bei den alternativ genannten Kriterien kann der Nutzungsgrad Hinweise zur Bedeutung einer Information liefern. Im Gegensatz zu den erstgenannten Kriterien lässt er sich aber ohne großen Aufwand ermitteln und automatisiert messen [Bh05], [Ch05], [CGY07], [Sha06]. Der Wert einer Information resultiert im Rahmen unseres Beitrags daher aus der Quantifizierung des Nutzungsgrades³. Dadurch wird eine berechenbare Grundlage zur anschließenden Klassifizierung der Informationen geliefert [Ch05], [Do04], [Sha06].

Eine Klasse bezeichnet das geeignete Speichermedium [MHP05], auf welches die Informationen verlagert werden sollen. Hierbei werden die Informationen in vier Klassen eingeteilt:

- Klasse „Online“: Informationen dieser Klasse werden oft von Anwendern für Analysen benötigt und sollten daher für den sofortigen und schnellen Zugriff [Pe05] im Online-Speicherbereich vorgehalten werden.
- Klasse „Nearline“: Informationen, auf die in regelmäßigen Abständen zugegriffen wird, für die aber eine Ablage auf den kostenintensiven Speichermedien des Data-Warehouse-Systems nicht lohnend ist, werden auf Nearline-Storage-Systemen [LLZ04], [MHP05] abgelegt.
- Klasse „Offline“: Informationen, auf die nur in Ausnahmefällen zugegriffen wird oder deren Aufbewahrung aufgrund rechtlicher Anforderungen notwendig ist, können auf kostengünstigen Speichermedien langfristig aufbewahrt werden [KSH96].⁴
- Klasse „Löschen“: Auf Informationen dieser Gruppe wird nicht mehr zugegriffen. Sie haben keinen Wert für das Unternehmen und können gelöscht werden, sofern keine rechtlichen Aufbewahrungsrestriktionen bestehen.

In Phase vier, „Informationsverlagerung“, werden die Informationen entsprechend ihrer Klasse auf die dafür vorgesehenen Speichermedien verlagert.

³ Vgl. Abschnitt 4.

⁴ Hierfür können Magnetbänder verwendet werden [Ja98]. Die Speicherung auf Magnetbändern gilt als kostengünstigste Alternative [HS96].

3 Anforderungen an Konzepte zur automatisierten Klassifizierung von Informationen

Mit Hilfe einer Literaturanalyse haben wir verschiedene Anforderungen an Konzepte zur automatisierten Informationsklassifizierung identifiziert. Diese werden im Folgenden vorgestellt.

Zunächst sollten Konzepte zur automatisierten Klassifizierung von Informationen die Kriterien Zugriffshäufigkeit und Zugriffszeitpunkt zur Bestimmung des Nutzungsgrades der Informationen verwenden. Ein Zugriff erfolgt, wenn eine Information gelesen oder verändert wird [Sha06]. Etwa sechzig bis achtzig Prozent der Informationen eines Unternehmens bleiben ungenutzt. Auf sie wird nur noch selten zugegriffen [Ch05], [Sha06], [Sh06], [Za04], was eine Auslagerung auf kostengünstigere Speicherbereiche rechtfertigt. Weiterhin muss ermittelt werden, zu welchen Zeitpunkten auf eine Information zugegriffen wurde, um zu bestimmen, ob Informationen vor kürzerer Zeit genutzt wurden oder ob für einen längeren Zeitraum nicht darauf zugegriffen worden ist [Ch05].⁵

Für die Automatisierung müssen darüber hinaus Funktionen implementiert werden, welche die Informationen anhand des Nutzungsgrades bewerten und anschließend den zugehörigen Klassen zuordnen [Ch05], [Me04], [Sha06], [Ve05]. Die Bewertungsfunktionen müssen in der Lage sein, basierend auf der Historie der Zugriffshäufigkeiten und -zeitpunkte, die zukünftigen Zugriffshäufigkeiten und -zeitpunkte zu prognostizieren bzw. zu berechnen [Ab06], [Do04], [Me04], [Sha06]. Damit wird die automatisierte Informationsklassifizierung vereinfacht, da Zugriffshäufigkeiten und -zeitpunkte sowie Klassenzugehörigkeiten nicht ständig neu berechnet werden müssen, sondern eine Vorausplanung bezüglich eines definierten Zeitraums erfolgen kann. Dies wirkt sich positiv auf die Performanz der Data-Warehouse-Systeme und damit auf die Kosteneinsparungspotentiale aus.

Trotz der Automatisierung sollte das Wissen der Anwender und Administratoren über die Systeme und Informationen berücksichtigt werden [Bh05], [Do04], [Pa05]. Anwender und Administratoren können wichtige Hinweise zur Verwendung der Informationen geben, beispielsweise auf Basis der aktuellen und zukünftigen Projekt- bzw. Auftragslage oder der Verwendung von Informationen innerhalb verschiedener Abteilungen des Unternehmens. Insbesondere bei der Definition der Bewertungsfunktionen und bei der Festlegung, ab welcher Zugriffshäufigkeit eine Information einer bestimmten Klasse angehören soll, ist dies von Relevanz. Außerdem ist das Wissen der Anwender und Administratoren für die Validierung der Klassifikation wichtig. Konzepte zur automatisierten Informationsklassifizierung sollten deshalb die Möglichkeit bieten, die Klassenzugehörigkeit auch nach der automatischen Klassifizierung manuell anzupassen. Dies würde es ermöglichen, nicht sinnvoll erscheinende Klassenzuordnungen noch vor der Informationsverlagerung zu revidieren. Dies ist insbesondere bei Vorschlägen zur Löschung von Informationen wichtig.

⁵ Vgl. Abschnitt 4.3. Dazu müssen durch das Data-Warehouse-System entsprechende Statistiken bzw. Metadaten protokolliert werden.

Neben dem Anwender- und Administratorenwissen sollte auch die Berücksichtigung rechtlicher Rahmenbedingungen ein Bestandteil der nachträglichen Validierung der Klassenzugehörigkeit sein, womit eine weitere wichtige Anforderung an Konzepte zur automatisierten Klassifizierung von Informationen identifiziert werden kann [MB07], [Sh06]. Muss eine Information aufgrund rechtlicher Restriktionen aufbewahrt werden, so darf sie nicht gelöscht werden, obwohl eventuell für einen längeren Zeitraum nicht auf sie zugegriffen wurde. Andererseits existieren Informationen, die nach einem gesetzlich vorgeschriebenen Zeitraum gelöscht werden müssen, obwohl oft darauf zugegriffen worden ist. In beiden Fällen ist eine nachträgliche Anpassung der Klassenzugehörigkeit der betreffenden Information nach einer automatisierten und auf Grundlage des Nutzungsgrades vorgenommenen Klassifikation notwendig.

Wesentliches Ziel von ILM ist die Einsparung von Kosten. Die automatisierte Informationsklassifizierung sollte daher nicht zuletzt zu einer Reduktion der Administrationskosten für Datenspeicher [Me04], [Sha06] führen. Weiterhin muss eine Reduktion der Aufbewahrungskosten aufgrund der Speicherung der Informationen auf dem jeweils optimalen Speichermedium erfolgen [Sha06], damit von einem effektiven ILM gesprochen werden kann.

Es sind bereits verschiedene Konzepte zur automatisierten Klassifizierung von Informationen publiziert worden. Verma et al. [Ve05] und Mesnier et al. [Me04] verwenden als Klassifikationskriterium den Typ der Information bzw. den Dateityp. Informationen gleichen Dateityps haben ihrer Meinung nach dieselben Eigenschaften und können einer gemeinsamen Klasse zugeordnet werden [Me04], [Ve05]. Zadok et al. [Za04] sowie Chandra, Gehani und Yu [CGY07] konzentrieren sich stark auf das Anwender- und Administratorenwissen über Data-Warehouse-Systeme und Informationen. Jeder Anwender und Administrator erzeugt eigene Bewertungsfunktionen bzw. bewertet und klassifiziert seine Informationen selbst. Shah et al. [Sha06] und Bhagwan et al. [Bh05] verwenden die Zugriffshäufigkeit und berücksichtigen zusätzlich das Anwender- und Administratorenwissen sowie die Performanz der Data-Warehouse-Systeme.

Chen [CH05] bezieht sich auf den Nutzungsgrad von Informationen und zieht zur Bewertung die Kriterien Zugriffshäufigkeit und Zugriffszeitpunkt heran. Bezüglich der Entwicklung unseres Klassifizierungskonzeptes orientieren wir uns an den Vorschlägen von Chen, welche im Folgenden vorgestellt werden. Im weiteren Verlauf des Beitrages werden diese Vorschläge auf den Data-Warehouse-Bereich übertragen und am Beispiel des SAP BI beispielhaft umgesetzt. Abschließend wird untersucht, inwieweit der von uns entwickelte Ansatz zur automatisierten Informationsklassifizierung in Data-Warehouse-Systemen die in diesem Abschnitt vorgestellten Anforderungen erfüllt.

4 Konzept zur automatisierten Klassifizierung von Informationen eines Data Warehouse

4.1 Bewertung von Informationen anhand des Verfahrens nach Chen

Das Verfahren nach Chen bestimmt einen Wert für Informationen zu einem bestimmten Zeitpunkt.⁶ Der ermittelte Wert basiert auf den Statistiken einer zuvor festgelegten Betrachtungsperiode. Diese Statistiken zeichnen auf, wann und wie oft ein Zugriff auf die betrachteten Informationen erfolgte. Daraus lassen sich Rückschlüsse über den Nutzungsgrad der Informationen ziehen.

Für die Bestimmung eines Wertes für Informationen auf Grundlage der Kombination von Zugriffshäufigkeit und Zugriffszeitpunkten teilt Chen die Betrachtungsperiode in Phasen gleicher Länge ein. Die Aufzeichnung der Statistiken erfolgt separat für jede Phase. Die Zugriffe auf Informationen innerhalb einer Phase sind gleichgewichtet. Für jede Phase wird ein Gewicht bestimmt. Je höher das Gewicht einer Phase ist, desto mehr trägt deren Zugriffsstatistik zur Bestimmung des Wertes einer Information bei. Hierdurch wird die Berücksichtigung der Zugriffszeitpunkte bei der Wertbestimmung erreicht. Die Summe der Gewichte aller Phasen der Betrachtungsperiode ist 1.

Die Länge der Phasen bestimmt, wie stark das Verfahren sich an den Zugriffszeitpunkten oder der Zugriffshäufigkeit ausrichtet. Ausgedehnte Phasen reduzieren den Effekt der Berücksichtigung von Zugriffszeitpunkten. Die Wertbestimmung erfolgt dann vorwiegend aufgrund der Zugriffshäufigkeit. Mit Hilfe einer Fallstudie hat Chen ermittelt, dass die Länge der Phasen bei einer Betrachtungsperiode von 60 Tagen zwischen 8 und 16 Tagen liegen sollte [Ch05]. In dieser Spanne ist das Verfahren robust gegenüber der Länge der Phasen.

Ein Zähler i verleiht jeder Phase einen Zahlenwert. Die Zählung beginnt in der Gegenwart. Weniger weit zurückliegende Phasen haben kleine Zahlenwerte, weit zurückliegende Phasen weisen höhere Zahlenwerte auf. Die Betrachtungsperiode (vp-valuation period) beginnt vor und endet mit dem Zeitpunkt t , zu dem der Wert der Informationen bestimmt werden soll. Die Anzahl der Phasen ist n , deren Länge in Tagen wird durch s bezeichnet.

$$vp = [t - (n \cdot s), t]$$

Unter der Bedingung, dass Informationen, auf die erst vor kurzem zugegriffen wurde, einen höheren Wert haben als Informationen, die vor längerer Zeit letztmalig genutzt worden sind, sollten die Gewichte der Phasen von der Vergangenheit zur Gegenwart hin steigen. Chen schlägt dazu folgende Berechnungsformel vor:

⁶ Vgl. Abschnitt 2, Vorgehensmodell Phase 3.

$$w_i = \frac{\left(\frac{1}{x}\right)^i}{\sum_{j=1}^n \left(\frac{1}{x}\right)^j} \quad \text{mit } x \geq 1$$

Hierbei stellt w_i das Gewicht der i -ten Phase dar. Der Parameter x ist frei wählbar. Dieser beeinflusst die Gewichtsverteilung der Phasen in der Betrachtungsperiode. Je größer x gewählt wird, desto steiler ist die Gewichtsverteilung. Chen betont, dass sehr flache oder sehr steile Gewichtsverteilungen vermieden werden sollten. Eine flache Gewichtsverteilung bei einem niedrigen x -Wert führt im Extremfall zur Gleichwertigkeit aller Phasen. Das Verfahren richtet sich dann vornehmlich an der Zugriffshäufigkeit aus. Eine besonders steile Gewichtsverteilung bei einem hohen Wert für x führt zur Vernachlässigung der Zugriffshäufigkeiten, da tendenziell alle Informationen, auf die in letzter Zeit zugegriffen worden ist, einen hohen Wert erhalten. Mit einem Wert für x zwischen zwei und drei ist das Verfahren laut Chen in der von ihr untersuchten Fallstudie hinreichend robust.

Der Wert einer Information ergibt sich aus der Multiplikation des Phasen-Gewichtes mit der Anzahl der Zugriffe innerhalb einer Phase, aufsummiert über alle Phasen der Betrachtungsperiode. Zur Anwendung dieses Bewertungsansatzes im Data-Warehouse-Bereich, schlagen wir folgendes Vorgehen vor.

4.2 Übertragung des Verfahrens nach Chen auf den Data-Warehouse-Bereich

Den Ausgangspunkt für die Klassifizierung von Informationen⁷ in einem Data Warehouse bilden multidimensionale Datenwürfel. Auf einem Datenwürfel sind Abfragen definiert, die von den Nutzern des Data-Warehouse-Systems ausgeführt und deren Ergebnisse für die Unterstützung von Entscheidungen herangezogen werden. Für jede Abfrage wird entsprechend des Ansatzes nach Chen ein Wert bestimmt. Dieser Wert wird auf alle Informationen (Datensätze) der Faktentabelle des Datenwürfels übertragen, die von der Abfrage angesprochen werden. Dafür wird unserem Klassifizierungskonzept eine relationale Datenbankarchitektur zugrunde gelegt. Im Ergebnis sind alle Informationen der Faktentabelle mit einem Wert hinterlegt, aufgrund dessen Rückschlüsse über den Nutzungsgrad der einzelnen Informationen gezogen und eine geeignete Klassifizierung vorgenommen werden können.

⁷ Im Kontext des Data Warehouse sind unter Informationen die in den multidimensionalen Datenwürfeln befindlichen Datensätze zu verstehen.

Die Beschränkung auf die Bewertung der Informationen aus der Faktentabelle eines Datenwürfels ist auf die Tatsache zurückzuführen, dass sich in den Faktentabellen der größte Teil des Informationsvolumens eines Data Warehouse befindet. Die Auslagerung von Informationen der Dimensionstabellen würde nicht nur zu Komplikationen bei der Durchführung von Auswertungen führen, sondern zudem keinen nennenswerten Beitrag zur Beschränkung des Informationsvolumens im Online-Speicherbereich im Rahmen eines effektiven ILM leisten.

Um den Wert für eine Abfrage d zum Zeitpunkt t ($v_t(d)$) zu bestimmen, wird nun analog zum in Abschnitt 4.1 beschriebenen Ansatz nach Chen das Gewicht einer jeden Phase mit der Anzahl der Ausführungen der Abfrage d innerhalb dieser Phase ($u_i(d)$) multipliziert und über die Betrachtungsperiode hinweg aufsummiert:

$$v_t(d) = \sum_{i=1}^n (w_i \cdot u_i(d))$$

Ein Beispiel soll die Berechnung des Wertes einer Abfrage verdeutlichen. Wir gehen dabei von einer Betrachtungsperiode mit drei Phasen aus. Abbildung 2 veranschaulicht die Nummerierung und Gewichtung der Phasen:

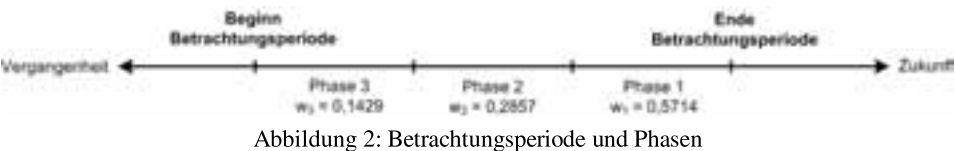


Abbildung 2: Betrachtungsperiode und Phasen

Für $x=2$ berechnet sich das Gewicht der ersten Phase wie folgt:

$$w_1 = \frac{\left(\frac{1}{2}\right)^1}{\sum_{j=1}^3 \left(\frac{1}{2}\right)^j} = \frac{0,5}{\left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3} = \frac{0,5}{0,875} = 0,5714$$

Auf dem von uns beispielhaft betrachteten Datenwürfel sind zwei Abfragen definiert. Tabelle 1 zeigt die Ausführungsstatistik für die zwei Abfragen in jeder Phase der Betrachtungsperiode.

	Phase 3	Phase 2	Phase 1
Abfrage A1	39	23	15
Abfrage A2	20	12	19

Tabelle 1: Ausführungsstatistik Beispiel

Der Wert für Abfrage A1 ergibt sich aus folgender Berechnung:

$$v_i(A1) = \sum_{i=1}^3 (w_i \cdot u_i(A1)) = 0,5714 \cdot 15 + 0,2857 \cdot 23 + 0,1429 \cdot 39 = 20,7152$$

Damit im Rahmen von ILM geeignete Regeln für die Klassifizierung der Informationen anhand ihres Nutzungsgrades aufgestellt werden können, sollten die Werte der Abfragen ein fest vorgegebenes Intervall (z. B. [0,10]) nicht verlassen. Um dies zu erreichen, müssen die Werte der Abfragen mit den vorzugebenden Intervallgrenzen normalisiert werden.

Nachdem für alle Abfragen des betrachteten Datenwürfels jeweils ein Wert bestimmt worden ist, wird dieser auf alle Informationen, die eine Abfrage angesprochen hat, übertragen. Werden Informationen von mehreren Abfragen angesprochen, ist den Informationen jeweils der höchste Wert zuzuordnen. Im Gegensatz zur normalen Durchschnittsbildung, bei der alle Abfragen gleichermaßen zur Bildung des Wertes für eine Information beitragen, kann diese Methode nicht zu einer Unterbewertung von Informationen führen. Eine gewichtete Durchschnittsbildung wäre eine Alternative dazu, ist aber aufgrund der zu bestimmenden Gewichte für die einzelnen Abfragen sehr viel schwerer automatisierbar.

Bei der Anwendung des vorgeschlagenen Bewertungsansatzes ist zu prüfen, ob innerhalb der eingegrenzten Betrachtungsperiode neue Abfragen entstanden sind oder gelöscht wurden. Treten solche Fälle auf, müssen die entsprechenden Abfragen gesondert behandelt werden, indem nur die Phasen zur Bestimmung des Wertes berücksichtigt werden, in denen die Abfrage tatsächlich existierte. Zudem sollten speziell für diese Abfragen neue Gewichte für die zu berücksichtigenden Phasen berechnet werden. Die neuen Gewichte müssen in ihrer Summe wieder 1 ergeben. Die Proportionalität zu den ursprünglichen Gewichten muss gewahrt bleiben. Um neben der Entstehung und Löschung von Abfragen auch Modifikationen mit dem vorgeschlagenen Bewertungsansatz abbilden zu können, ist die Einführung einer Versionsverwaltung für Abfragen grundsätzliche Voraussetzung. Es muss jederzeit bestimmbar sein, welche Informationen eine Abfrage zu einem bestimmten Zeitpunkt angesprochen hat. Die Behandlung von Abfragen, deren Version sich innerhalb der Betrachtungsperiode geändert hat, kann dann analog zur Berücksichtigung von neu entstandenen und gelöschten Abfragen erfolgen, indem eine neue Version wie eine eigenständige Abfrage behandelt wird.

Im Anschluss an die Bewertung der Informationen werden diese entsprechend vordefinierter Regeln in verschiedene Klassen eingeteilt. Diese Regeln sind unter Einbeziehung des Anwender- und Administratorenwissens über das betrachtete Data-Warehouse im Vorfeld der Klassifizierung zu definieren und legen fest, welche Werte der Informationen zu welcher Klasse gehören. Auf Grundlage dieser Klassifizierung der Informationen kann die Verlagerung auf die einzelnen Speicherbereiche erfolgen.⁸

⁸ Vgl. Abbildung 1, Phase 4.

4.3 Automatisierte Klassifizierung von Informationen am Beispiel des SAP BI

Die praktische Anwendbarkeit des Konzeptes wurde im Rahmen eines SAP-BI⁹-Projektes überprüft. Dieses Projekt wurde im Jahr 2007 im Rahmen einer Diplomarbeit in Zusammenarbeit mit dem Praxispartner IBM Deutschland in Chemnitz realisiert. Die Überprüfung erfolgte in Form einer prototypischen Implementierung des Bewertungsverfahrens einschließlich einer Klassifizierung der Informationen anhand vordefinierter Regeln.

Eine wesentliche Voraussetzung der Anwendung des Konzeptes besteht darin, dass durch das SAP BI statistische Informationen über die Ausführung von Abfragen aufgezeichnet werden. Das System ist in der Lage zu protokollieren, welche Abfrage zu welchem Zeitpunkt ausgeführt worden ist.¹⁰ Zur Implementierung des vorgeschlagenen Bewertungskonzeptes musste aus den aufgezeichneten Informationen des SAP BI eine geeignete Ausführungsstatistik erzeugt werden. Dabei ermittelt der implementierte Prototyp die Summe der Ausführungen einer Abfrage innerhalb einer Phase, wofür Beginn und Ende der Betrachtungsperiode sowie die Anzahl bzw. Länge der Phasen im Vorfeld festzulegen sind. Des Weiteren muss der Parameter x vorgegeben werden, um die Phasen für die folgende Wertbestimmung entsprechend gewichten zu können.

Nach der Bewertung aller Abfragen eines Datenwürfels und der Normalisierung der Abfrage-Werte im Intervall $[0,10]$ werden im Rahmen der prototypischen Implementierung des Konzeptes die Abfrage-Werte auf alle Informationen der Faktentabelle übertragen, die jeweils von einer Abfrage angesprochen werden. Es kommt die beschriebene Maximum-Methode zum Einsatz. Um die von einer Abfrage bei deren Ausführung angesprochenen Informationen der Faktentabelle zu bestimmen, wurden bei der Implementierung des Verfahrens die Selektionsbedingungen der Abfrage ausgewertet.

Im Anschluss an die Bewertung der Informationen werden diese entsprechend vordefinierter Regeln in verschiedene Klassen eingeteilt. Bei der Implementierung des Prototyps wurde die in Tabelle 2 dargestellte Klassifizierung gewählt¹¹:

Wert der Information	Klasse
10 – 7	Online
6 – 4	Nearline
3 – 1	Offline
0	Löschen

Tabelle 2: Beispielhafte Klassifizierung

⁹ Hierbei handelt es sich um das von der SAP AG vertriebene Produkt SAP NetWeaver Business Intelligence in der Version 7.0.

¹⁰ Hierfür stellt die SAP AG vorgefertigte Objekte innerhalb des so genannten SAP Business Intelligence Content zur Verfügung.

¹¹ Vgl. die unterschiedlichen Klassen, Abbildung 1, Phase 4.

4.4 Auswertung der Klassifizierungsergebnisse

Untersuchungen auf einem künstlich erzeugten Informationsbestand ergaben, dass durch die Implementierung des vorgeschlagenen Klassifizierungskonzeptes bis zu fünfzig Prozent des Informationsbestandes eines Datenwürfels aus dem Online-Speicherbereich ausgelagert werden kann. Grundlage des künstlich erzeugten Informationsbestandes ist ein Datenwürfel mit 50.095 Informationen. Bei diesen Informationen handelt es sich um anonymisierte Finanzinformationen eines Unternehmens, die von dem Praxispartner IBM Deutschland für die Zwecke der Diplomarbeit zur Verfügung gestellt wurden. Auf dem untersuchten Datenwürfel wurden elf Abfragen mit willkürlich festgelegten Selektionsbedingungen angelegt. Diese Abfragen wurden nach einem definierten Ausführungsplan auf den Datenwürfel angewandt, um möglichst unterschiedliche Ausführungsmuster zu erzeugen und deren Auswirkungen auf die Bewertung der Informationen abschätzen zu können. Untersuchungen mit den von Chen vorgeschlagenen Standardwerten für den Parameter x und die Anzahl bzw. Länge der Phasen ergaben, dass etwa ein Drittel der Informationen der Faktentabelle in den Nearline-Bereich und nahezu ein Sechstel in den Offline-Bereich verschoben werden können. Weniger als ein Prozent der Informationen wurden zur Löschung vorgeschlagen.

Weiterführende Untersuchungen bezüglich der Eingabewerte für den Parameter x und der Anzahl bzw. Länge der Phasen ergaben, dass bei einer Verringerung des Standardwertes für x von 2 auf 1,2 die Abfragen tendenziell höher bewertet werden. Abbildung 3 verdeutlicht dies.

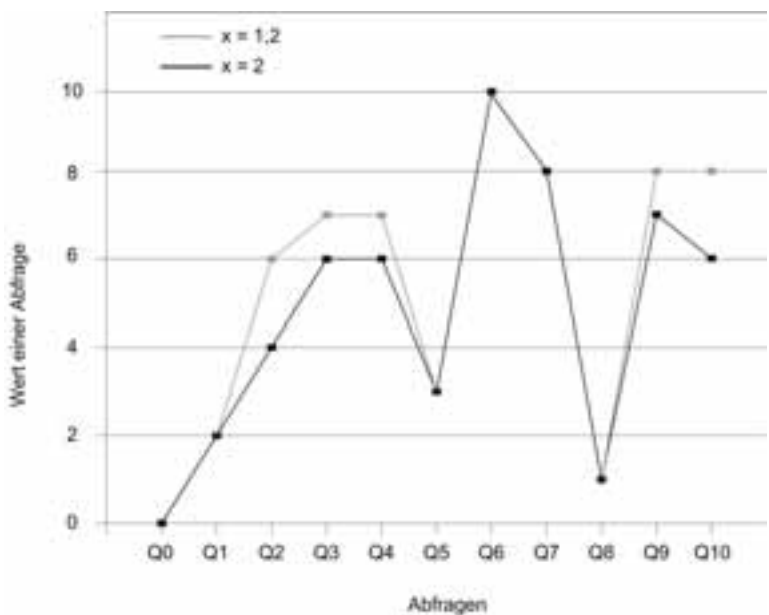


Abbildung 3: Steigende Abfrage-Werte bei sinkendem Parameter x

Der Grund dafür liegt darin, dass sich mit sinkendem Parameter x die Gewichte der Phasen gegenseitig annähern und der Einfluss der Phasen auf die Bewertung der Abfragen sinkt. Die Abfragen werden in diesem Fall tendenziell stärker an der Häufigkeit ihrer Ausführung als an den Ausführungszeitpunkten bewertet, da die sehr flache Gewichtsverteilung der Phasen nicht mehr für eine Abwertung der Ausführungen aus weiter zurückliegenden Phasen sorgt. Dies führt bei dem gegebenen Informationsbestand dazu, dass nur noch ein Sechstel der Informationen der Faktentabelle aus dem Online-Speicherbereich zur Auslagerung vorgeschlagen wird. Je höher der Parameter x ist, desto mehr werden die Ausführungshäufigkeiten bei der Bewertung der Abfragen vernachlässigt. Die Bewertung erfolgt dann tendenziell nach dem Zugriffszeitpunkt.

Gleiche Untersuchungsergebnisse ergaben sich bei Veränderung der Anzahl bzw. Länge der Phasen. Bei der Bewertung der Abfragen werden die Ausführungszeitpunkte tendenziell vernachlässigt, sobald bei einer Betrachtungsperiode von acht Wochen die Anzahl der Phasen kleiner als fünf ist. Wie bereits von Chen beschrieben, ist der Parameter x und die Anzahl bzw. Länge der Phasen so zu wählen, dass Ausführungshäufigkeit und -zeitpunkte gleichermaßen an der Bestimmung der Abfrage-Werte beteiligt sind.

4.5 Evaluierung des Klassifizierungskonzeptes

Es soll nun überprüft werden, inwiefern das von uns entwickelte Konzept die in Abschnitt 3 definierten Anforderungen erfüllt.

Die Kriterien Zugriffshäufigkeit und -zeitpunkte bei der Bestimmung des Nutzungsgrades wurden analog zum Ansatz von Chen verwendet. Damit kann überprüft werden, wie oft und wann auf die zu einem Datenwürfel erstellten Abfragen zugegriffen wird bzw. welche Informationen gelesen werden.

Mit der automatisierten Bewertung und Klassifizierung der Informationen ist eine weitere Anforderung erfüllt. Die Berücksichtigung des Anwender- und Administratorenwissens erfolgt bei der Bestimmung der Betrachtungsperiode, der Anzahl bzw. Länge der Phasen, der Festlegung des Parameters x sowie bei der Aufstellung der Regeln für die Klassifizierung der Informationen. Durch diese Möglichkeiten der Einflussnahme in den Prozess der Informationsklassifizierung ist das Verfahren an die Nutzungsintensität des betrachteten Data-Warehouse-Systems flexibel anpassbar. Eine manuelle Veränderung der Klassifizierung im Anschluss an die automatisierte Klassifikation ist in unserem Konzept nicht möglich.

Die Prognose zur Ermittlung zukünftiger Zugriffshäufigkeiten und -zeitpunkte sowie die Anforderung zur Berücksichtigung rechtlicher Restriktionen erfüllt unser Konzept nicht. Die Informationen werden ausschließlich anhand ihrer Nutzung auf Grundlage von Vergangenheitswerten bewertet. Neben unserem Konzept erfüllt auch keines der in Abschnitt 3 vorgestellten Konzepte zur automatisierten Klassifizierung von Informationen die Berücksichtigung rechtlicher Anforderungen. Dieses Gebiet ist ein offenes Forschungsthema [Ab06], [MB07].

Die Berücksichtigung des zukünftigen Nutzungsgrades ließe sich in unser Bewertungsverfahren integrieren, indem die Betrachtungsperiode, die der Bewertung zugrunde liegt, in die Zukunft ausgedehnt wird. Chen schlägt vor, dass die Gewichtung der Phasen dann in Form einer Glockenkurve erfolgt, die ihren Höhepunkt zur gegenwärtigen Phase erreicht. Zukünftige Phasen erhalten analog zur Vergangenheit sinkende Gewichte, da der Nutzungsgrad der Informationen nur geschätzt und daher unsicher ist. Zur Konkretisierung dieses Vorgehens und dessen Integration in unser Klassifizierungskonzept ist weitere Forschungs- und Entwicklungsarbeit notwendig.

Ein wesentlicher Kritikpunkt an dem von uns vorgeschlagenen Klassifizierungskonzept ist, dass Informationen, die in den Offline-Bereich verlagert wurden, nicht mehr direkt auswertbar sind. Sobald Informationen aber aus dem Bereich der direkten Auswertbarkeit ausgelagert werden, ist die Interpretation der verbleibenden Informationen ohne entsprechende Kennzeichnung sehr schwierig, da beispielsweise Aggregationen nicht mehr vollständig sind. Wie eine solche Kennzeichnung genau aussehen kann, ist ein weiterer Punkt, auf den im Rahmen weiterer Arbeiten näher eingegangen werden muss [St01].

Unser Klassifizierungskonzept behandelt ausschließlich die Klassifikation von Informationen in Datenwürfeln, nicht aber in physischen Aggregaten. Die Berücksichtigung von Aggregaten ist nicht notwendig, da diese ohnehin einer statistischen Auswertung durch das eingesetzte Data-Warehouse-System unterliegen und das System selbst Vorschläge zur Verwendung der in den Aggregaten befindlichen Informationen unterbreitet.¹² Aggregate ermöglichen einen schnellen Zugriff auf die Informationen eines Datenwürfels. Dabei werden die Informationen des Datenwürfels, auf welche innerhalb von Berichten häufig zugegriffen wird, redundant gespeichert, um eine höhere Performanz zu erzielen. Aus diesem Grund sollten die Informationen der Aggregate ausschließlich innerhalb der Speicherkategorie „Online“ vorgehalten werden.¹³ Werden Informationen in Datenwürfeln verändert, so werden diese durch das so genannte „Hochrollen“ der Informationen in die zugehörigen Aggregate übertragen. Damit wirken sich Auslagerungen von Informationen aus den Datenwürfeln auch automatisch auf alle damit verbundenen physischen Aggregate aus.

Das Potential unseres Klassifizierungskonzeptes wurde seitens des Praxispartners IBM Deutschland erkannt und wertgeschätzt, so dass derzeit an der Weiterentwicklung der prototypischen Realisierung in SAP BI gearbeitet wird. Als hauptsächliche Herausforderung stellt sich hierbei der Ausbau des Vorgehens zur Bestimmung der von einer Abfrage angesprochenen Informationen. Neben statisch definierten Selektionsbedingungen sollen nun auch variable, während der Abfrageausführung festgelegte Selektionsbedingungen ausgewertet werden. Das sich hierbei ergebende Problem besteht darin, dass in SAP BI die dynamischen Selektionsbedingungen nur protokolliert werden, sofern zur Ausführung der Abfrage auf die Datenbank zurückgegriffen werden muss. Erfolgt das

¹² Detaillierte Aussagen zur Verwendung bzw. zu den Zugriffen auf die Informationen von Aggregaten beispielsweise in SAP BI liefert die BI-Statistik.

¹³ Vgl. Abschnitt 2

Lesen der Abfrageergebnisse jedoch vollständig aus dem Zwischenspeicher des Systems, findet keine Aufzeichnung der verwendeten Selektionsbedingungen statt. Hier weist das SAP BI bezüglich der Realisierung unseres Konzeptes ein klares Defizit auf.

5 Zusammenfassung und Ausblick

Unser Beitrag konzentriert sich auf die zentrale Teilaufgabe des ILM, die Informationsklassifizierung. Nach der Einordnung dieser Teilaufgabe in den Gesamtkontext des ILM im Data-Warehouse-Bereich haben wir auf Grundlage der Vorarbeiten von Chen ein Konzept vorgestellt, welches Informationen in einem Data Warehouse entsprechend ihres Nutzungsgrades automatisiert klassifiziert. Anhand einer prototypischen Implementierung im Rahmen eines SAP-BI-Projektes konnten wir die Praktikabilität des Konzeptes demonstrieren und Verbesserungsmöglichkeiten aufzeigen.

Der Vorteil unseres Klassifikationskonzeptes besteht in der automatisierten Ermittlung der Zugriffshäufigkeiten auf einzelne Informationen in Datenwürfeln. Dieses Verfahren wurde speziell für das Data-Warehouse-Umfeld entwickelt, wodurch unser Konzept nicht ohne Weiteres auf OLTP-Systeme übertragen werden kann.

Im Fokus der zukünftigen Forschung und Entwicklung steht die Weiterentwicklung von ILM-Software [GM05]. Eine Software, die automatisiert Vorschläge bezüglich des Wertes bzw. der Klassifikation von Informationen liefert, ist für die Entwicklung und Umsetzung von ILM-Konzepten im Data-Warehouse-Umfeld von Vorteil, wenngleich Einriffsmöglichkeiten der Verantwortlichen bzw. Anwender weiterhin notwendig sind.

6 Literaturverzeichnis

- [Ab05] Abd-El-Malek, M. et al.: *Ursa Minor: versatile cluster-based storage*. In (Gibson G. Hrsg.): *Proceedings of the 4th USENIX Conference on File and Storage Technologies*, San Francisco, 2005; S. 59-72.
- [Ab06] Abd-El-Malek, M. et al.: *Early experiences on the journey towards self-* storage*. In (Lomet, B. Hrsg.): *Data Engineering Bulletin*, Ausgabe 29, Nr. 4, Atlanta, 2006; S. 55-62.
- [Al01] Allen, N.: *Don't waste your storage dollars: what you need to know*. Research Note, Gartner Group Inc., Stamford 2001.
- [Bh05] Bhagwan R. et al.: *Time-varying Management of Data Storage*. In (Candea G.; Openheimer, D. Hrsg.): *First Workshop on Hot Topics in Systems Dependability*, Yokohama, 2005.
- [CGH05] Chamoni, P.; Gluchowski, P.; Hahne, M.: *Business Information Warehouse: Perspektiven betrieblicher Informationsversorgung und Entscheidungsunterstützung auf der Basis von SAP Systemen*. Berlin, Heidelberg, New York, 2005; S. 36-39.
- [CGY07] Chandra, S.; Gehani, A.; Yu, X.: *Automated Storage Reclamation Using Temporal Importance Annotations*. In: *27th International Conference on Distributed Computing Systems*, Toronto, 2007; S. 12.

- [Ch05] Chen Y.: Information valuation for Information Lifecycle Management. In (Parashar, M. et al. Hrsg.): Proceedings of the 2nd International Conference on Autonomic Computing, Seattle, Washington, 2005; S. 135-146.
- [Do04] Douglass, F. et al.: Position: Short Object Lifetimes Require a Delete-Optimized Storage System. In: Proceedings of the 11th workshop on ACM SIGOPS European workshop: beyond the PC, ACM Press, New York, 2004.
- [El03] Ellard, D. et al.: Attribute-Based File Prediction of File Properties. Harvard, 2003; S. 1-14.
- [GM05] Gillet, F.; Mendel, T.: Organic IT: IT-Kosten senken, Unternehmensabläufe beschleunigen. In: (Kuhlin, B.; Thielmann, H. Hrsg.): Real-time Enterprise in der Praxis: Fakten und Praxis. Berlin, Heidelberg, New York, 2005; S. 483-502.
- [Gr03] Gray, J.: A conversation with Jim Gray. In (ACM Press Hrsg.): Queue Storage, Ausgabe 1, Nr. 4, New York, 2003; S. 8-17.
- [HS96] Hillyer, B.; Silberschatz, A.: Random I/O scheduling in online tertiary storage systems. In: ACM SIGMOD International Conference on Management of Data, 1996; S. 195-204.
- [IIS01] Inmon, W.; Imhoff, C.; Sousa, R.: Corporate Information Factory. 2. Auflage, New York et al. 2001; S. 139-156.
- [Ja98] Janko, W.: Informationswirtschaft 1. Grundlagen der Informatik für die Informationswirtschaft. Springer-Verlag, Berlin, 1998.
- [KSH96] Küspert, K.; Schaarschmidt, R.; Herbst, A.: Archiv Transaktionen: Ein Ansatz für asynchrones, transaktionsgesichertes Archivieren großer Datenmengen in Datenbanksystemen. In: ITG-Fachbericht zur 4. ITG/GI/GMA-Fachtagung Softwaretechnik in Automation und Kommunikation – Rechnergestützte Teamarbeit (STAK). München, 1996; S. 195-211.
- [LLZ04] Liu, B.; Li, J.; Zhang, Y.: Optimal Data Dispatching Methods in Near-Line Tertiary Storage System. In: (Li, Q.; Wang, G.; Feng L. Hrsg.): Proceedings of the 5th International Conference: Advances in Web-Age Information Management. Dalian, China, 2004, S. 690-695.
- [MB07] Mont, M.; Beato, F.: On Parametric Obligation Policies: Enabling Privacy-aware Information Lifecycle Management in Enterprises. In (IEEE Computer Society): Proceedings of the 8th IEEE International Workshop on Policies for Distributed Systems and Networks, Washington, 2007; S. 51-55.
- [Me04] Mesnier, M. et al.: File classification in self-* storage systems. In (Ibrahim, M. Hrsg.): Proceedings of the 1st International Conference on Autonomic Computing (ICAC-04). New York 2004; S. 44-51.
- [MHP05] Maier, R.; Hädrich, T.; Peinl, R.: Enterprise Knowledge Infrastructures. Springer-Verlag, Berlin, Heidelberg, New York 2005; S. 109, 116, 254-257.
- [Mo06] Mont, M.: On Privacy-aware Information Lifecycle Management in Enterprises: Setting the Context. In (Paulus, S.; Pohlmann, N.; Reimer, H. Hrsg.): ISSE 2006 – Securing Electronic Business Processes. Vieweg-Verlag, Wiesbaden 2006; S. 405-415.
- [MS08] Matthesius, M.; Stelzer, D.: Analyse und Vergleich von Konzepten zur automatisierten Informationsbewertung im Information Lifecycle Management. In (Bichler, K. et al. Hrsg.): Multikonferenz Wirtschaftsinformatik 2008. München, 2008; S. 471-482.

- [MW99] Moody, D.; Walsh, P.: Measuring the value of information: An asset valuation approach. In (Heje, J. P. et al. Hrsg.): Proceedings of the 7th European Conference on Information Systems. Copenhagen, Frederiksberg, 1999; S. 496-512.
- [Pa05] Patrascu, R. et al.: New approaches to optimization and utility elicitation in autonomic computing. Technical Report, Toronto, 2005.
- [Pe05] Petrocelli, T.: Data Protection and Information Lifecycle Management. Prentice Hall International, New York et al., 2005; S. 23.
- [Sha06] Shah, G. et al.: ACE: Classification for Information Lifecycle Management. Almaden et al., 2006; S. 1-15.
- [Sh06] Short, J.: Information Lifecycle Management: An Analysis of End User Perspectives. San Diego, 2006; S. 2-37.
- [St01] Stock, S.: Modellierung zeitbezogener Daten im Data Warehouse. Zugleich Dissertation an der Universität Duisburg 2000, Wiesbaden 2001; S. 133.
- [TS07] Thome, G.; Sollbach W.: Grundlagen und Modelle des Information Lifecycle Management. Springer-Verlag, Berlin, Heidelberg, New York, 2007; S. 22-30, 152, 171-172.
- [Ve05] Verma, A. et al.: An Architecture for Lifecycle Management in Very Large File Systems. In: 13th NASA Goddard, 22nd IEEE Conference on Mass Storage Systems and Technologies (MSST 2005), Monterey, 2005; S. 160-168.
- [WW07] Wixom, B.; Watson, H.: Introduction to the Minitrack on Data Warehousing and Business Intelligence. In: Proceedings of the 40th Hawaii International Conference on System Sciences, 2007; S. 214.
- [YB02] Yates-Mercer, P.; Bawden, D.: Managing the paradox: the valuation of knowledge and knowledge management. In: Journal of Information Science, Ausgabe 28, Nr.1, London, 2002; S. 19-29.
- [Za04] Zadok, E. et al.: Reducing Storage Management Costs via Informed User-Based Policies. In: 12th NASA Goddard, 21st IEEE Conference on Mass Storage Systems and Technologies (MSST 2004), Maryland, 2004; S. 101-105.

Harmonizing company-wide Information Objects

Alexander Schmidt, Boris Otto

Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedberg-Str. 8
9000 St. Gallen
{alexander.schmidt | boris.otto}@unisg.ch

Abstract: In today's companies, particularly multi-national enterprises acting on a global scale, historically grown systems and application landscapes, as well as processes, lead to significant problems when systems (and even employees) need to communicate with each other, i.e. when entities are exchanged across processes and organizational units. What is needed is a common language and, hence, unambiguously and consistently defined entities that represent essential objects of a company's environment. The paper at hand introduces a method that is intended to enable companies to increase the transparency on and consistency among information objects on a conceptual level. The scope of METIO is not limited to a one-time effort, but rather constitutes an iterative approach for a continuous perpetuation and improvement of a consistent set of information objects. The whole method is designed for application in the context of large-scale companies that, due to their size and international scope, dispose of a certain complexity and inconsistency in regard to their information objects.

1 Introduction

In today's companies, particularly multi-national enterprises acting on a global scale, historically grown systems and application landscapes, as well as processes that are not harmonized and consistent, are nothing unusual. Problems arise when systems (and even employees) need to communicate with each other, i.e. when entities are exchanged across processes and organizational units. What is needed is a common language and, hence, unambiguously and consistently defined entities that represent essential objects of a company's environment [NL06]. "A key challenge of data quality is an incomplete or unclear set of semantic definitions of what the data is supposed to represent, in what form" [BeDu07].

The consequences of data that is or can be misunderstood may be the malfunction of business processes, applications and whole organizations. In its worst case, mission-critical decisions can be based on wrong information because of false data interpretation. Examples of disastrous implications are manifold, be it the destruction of the Mars Climate Orbiter due to the usage of different metrics [Ne05a], the death of patients because of prescription mistakes and mislabelled blood samples or the additional costs of the United States Postal Service resulting from undeliverable mail [Pi05]. These examples

illustrate the importance of information as a company asset which should be managed actively with an organization-wide, architected approach. Data must be shareable in a transparent fashion across the organization and be under corporate ownership, rather than serving personal interests [To99].

The concept of (business) metadata provides a promising approach to reduce these misunderstandings by adding context to data [Ne05a]. It aims at making meaning explicit and providing definitions to important business terms and objects, data elements and abbreviations. However, this approach raises new questions: How do I determine which objects to focus on? How do I store and manage metadata? What roles do I need to define in order to embed the concept within my organization? How do I guarantee that everybody can have constant access to them during his operational work?

Most of the concepts dealing with metadata are rather vague and lack a guided procedure particularly for keeping data quality high. Moreover, they are usually not aligned with the business needs and the actual users (due to an overload of technical information) leading to insufficient use [OI03]. The paper at hand provides a methodology that addresses this topic and gives answers to the questions mentioned above. It shows how metadata repositories (such as a Business Data Dictionaries) can be employed for managing data quality and presents a method – that we call METIO (Method for Establishing Transparency on Information Objects) – that enables an organization to successfully create and, most notably, keep transparency on and consistency among business relevant information objects (IO) with the help of metadata.

The remainder of the paper is structured as follows: Section 2 gives an overview on the action research approach pursued and the projects which form the collaborative environment for the development of METIO. The third chapter provides the conceptual background for our research by firstly outlining a synoptic definition for metadata based on respective literature and thereafter examining the elements necessary to describe a comprehensive method in our research field. Chapter 4 contains the METIO methodology in more detail and, based on a peer literature review, derives how information objects can be unambiguously defined. The paper closes with a short conclusion and the outlook for further research.

2 Research Approach

The research follows the principles of design science which state that useful solutions must be obtained through the design and evaluation of models, methodologies and systems [WH07]. The research provides both solutions relevant for defined business requirements and contributes to the advancement of the scientific body of knowledge [He04]. The context of this research is set by a university research project, which is being conducted in collaboration with industrial partner companies. Development of the proposed method METIO involved both university researchers and subject matter experts from the partner companies. In the course of the collaboration, an action research approach was followed in order to provide the researchers with a detailed and continuous flow of information on the research subject and its context [CH98]. It accounts for the

fact that complex social systems in which humans interact using information technologies can only be understood by inducing change and observing the effects of such change ([BP99], [Ni08]). The iterative five step approach recommended by [BP99] consisting of 1) Diagnosing, 2) Action planning, 3) Action taking, 4) Evaluating, and 5) Specifying learning, serves as the guideline for our research and is illustrated for our specific case in Figure 1.

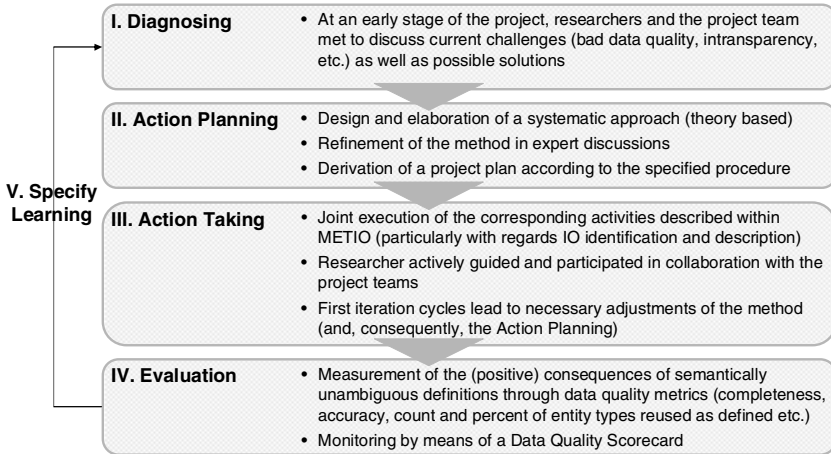


Figure 1: Pursued research process

The two real-life cases in which we conduct our action research are shortly introduced in the following. Their contribution to our research was twofold: firstly, they provided valuable input for the refinement of the theoretically derived procedure model (see Chapter 4.1) as well as the identification of attributes necessary to unambiguously describe an information object (see Chapter 4.2) during the Action Planning phase. Secondly, the cases enabled a real-world application of METIO, allowing for a comprehensive evaluation of the method and, consequently, significant findings for further improvement.

2.1 Public Infrastructure Operator

The company is a national European railway network operator. Being part of a state-owned corporation, its main business objectives are the provision of non-discriminating access to the railway infrastructure, the development of customer-oriented route offerings, and the maintenance, operation and advancement of the railway network.

The company is facing compliance requirements (e.g. issued by the European Commission or the national network agency), increasing reporting needs (e.g. infrastructure registry for shareholders), and cost reduction initiatives. Among the most important weaknesses in data management, which hinder the fulfilment of these requirements, are the following:

- Lack of responsibilities for data and information objects leading to ambiguous definitions and understandings of the commonly used objects;
- Operational data is inconsistent which causes poor data quality in analytical systems such as data warehouses (which form the basis for the infrastructure registry);
- Data models are system-specific; no integrated view exists across system boundaries;
- Data flows between different systems are neither transparent nor comprehensively managed so that process owners do not know what the original data sources are which they rely on within their processes and processes themselves suffer from poor interfaces with manual interaction.

At present, the company is undertaking an initiative to establish a common infrastructure data management aiming at the increase of transparency of the fundamental information objects and at efficiency gains in the data management domain.

2.2 Automotive Manufacturer

The case refers to the passenger cars division of an international automotive manufacturer. The division operates ten major plants across the globe. The business factors with impact on data management mainly derive from the overall corporate profitability targets. They materialize in the need for common reporting structures to allow for comparison of different locations and in the constant requirement to reduce general and administrative costs.

Among the reasons for current shortcomings in data management meeting the requirements is the complexity of the application architecture. In the past, application planning was decentralized, i.e. various business units were responsible for the task, leading to a total number of more than 2,000 applications in operation. Today, with integrated business processes that span multiple units, difficulties occur with mismatching definitions of information objects, unclear source systems for certain data objects, and numerous point-to-point connections between different application systems.

Apart from that, with continuing demands to reduce general and administrative costs, also the cost of IT has to be reduced. However, with the current lack of transparency regarding the relationship of information objects to business processes as well as application systems, consolidation of the system landscape is not an option as it is not clear which systems hold redundant data or data critical to the business, and which systems serve which business processes.

As a response to that, part of the architecture planning initiative is an effort to create transparency regarding information objects, especially with regard to their relation to business objects used in business processes and to data objects held in application systems.

3 Conceptual Foundation

3.1 Master data

In literature data are most commonly differentiated due to their purpose of use and their frequency of modification. Based on these dimensions, we can distinguish between master, inventory, change and transaction data ([We01], [HN05]). We deliberately constrain the application of METIO on master data. The reason for concentrating on master data can be seen in their characteristics: Master data represent an organization's core data entities [Wh06] that are rarely subject to change. They are used company-wide across different business processes and by multiple application systems. Master data are, therefore, specified by a relatively large number of attributes, often referenced by transaction data. Their enterprise-wide usage leads to increased complexity as master data are accessed by a multitude of employees. This shows their importance and value for companies and leads to an increased need for proper master data definition.

3.2 Definition and Significance of Metadata

Metadata can be defined in general as data that describes other data (their meaning and properties) [Bu99] demarcating them from other data types, such as transaction and master data. More precisely, we use the term to determine important characteristics that need to be known for either database and application engineering [En99] or the general, semantically unambiguous understanding of data within the enterprise. Metadata, accordingly, facilitates the identification, retrieval, use and management of data as they allow an organization to better understand its data sources and definitions [MS06]. TOZER summarizes their function by "seeing metadata as the means by which the structure and behavior of data is recorded, controlled, and published across an organization" [To99]. The most comprehensive definition is provided by MARCO who describes metadata as "all physical data (contained in software) and knowledge (contained in employees) from inside and outside an organization, including information about the physical data, technical and business processes, rules and constraints of the data, and structures of the data used by a corporation" [Ma00]. The significance of this definition derives from its strong business orientation that we pursue in this paper as well. Herein, we utilize metadata in the form of attributes (see Chapter 4.2) that need to be defined for specifying syntax and semantics of each information object.

Due to their high semantic content, metadata are the fundamental components for the design of information object models, as well as Business Data Dictionaries (BDD), serving as an original source for the definition of data elements and possess, hence, a high strategic value for companies [Ch06]. The strategic value primarily results from the significance the use of metadata has for evaluation and improvement of data quality by unambiguously characterizing information objects. By maintaining information about the source of data, their (change) history or responsibility, metadata facilitates the challenge of keeping data consistent, accurate and complete. And high quality data, in turn, is pivotal for enabling service-oriented business applications [NL06], for helping to increase the validity of strategic decisions [SZW03] and allowing high regulatory com-

pliance [Fr06]. Moreover, they enable a time- and cost-efficient way of retrieving, managing, evaluating and using appropriate information through precise queries which increases the confidence of users in data and augments the decision-making quality [Ma00]. The semantic content is essentially provided by standard (textual) definitions of the according data entities. Metadata management denotes the assignment of these definitions to data as well as their maintenance in a centralized metadata repository, such as a BDD [DL06].

3.3 Distinction between Business, Information and Data Objects

Referring to information objects, we would like to clearly demarcate the concept at this point of the paper from related terms, such as business objects and data objects. Within this paper we position business objects on a process level representing the input and output of business tasks, e.g. the entities that are exchanged within and between business processes. These business objects are relevant to business experts and generally described (if at all) in a simple textual form or an enumeration of their constituent attributes – similar to the business object description proposed by [Sc01]. The definition contains a coarse-granular description of characteristics relevant from a business perspective. Data objects on the other hand, are technical representations of these business objects on a system level. In most cases these entities are mapped in a more formalized way and contain more technical attributes, such as data types, field lengths etc. In between, we include an additional level containing information objects that constitute business relevant entities on a logical level. Information objects are described with their semantics as well as structure (consisting of relationships to other information objects) and, consequently, go beyond purely business-oriented definitions. They are mapped and described with their entire set of attributes and consequently represent an integrated, cross-applicational view of both business and data object characteristics. By contrast, data objects are application-specific storing a subset of characteristics of the corresponding information object [Sc05].

The tripartite differentiation corresponds to general approaches such as the Object Management Group's Model Driven Architecture (MDA). Within this framework the OMG distinguishes between three different viewpoints and, hence, models. In conformity with our three-level structure, the Computation Independent View focuses on the environment of a system (processes in our case) independently from the concrete implementation with the domain practitioner as the central target group [OM03]. This corresponds to our business objects. On the level below, the Platform Independent Model, the system is still considered on a conceptual level, unaffected by the underlying platform, but already specifying concrete operations of the system. And finally, the Platform Specific View characterizes in detail how a system uses a particular type of platform [OM03].

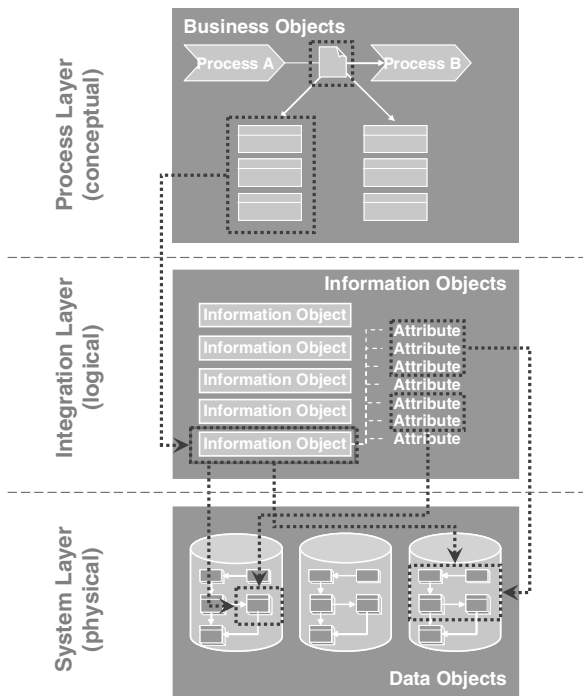


Figure 2: Three-layer differentiation between business, information and data objects

Finally, we would like to point out that we restrain our remarks within this paper on a type-level only and do not consider instances of the three object types.

4 A Method for Establishing Transparency on Information Objects

As defined in literature ([He93], [Gu94]) a method needs to contain more than just a procedure model and above all establish a metamodel and define roles that are responsible for carrying out each of the specified activities. Particularly the latter are very specific in the field of Data Quality Management and not intuitively comprehensible. Due to space limitations, we will not specify these parts of the method in this paper. However, we refer to well-established role definitions, as can be found in respective Data Governance literature ([We07], [En99], [DL06]) and use the roles of Chief, Business and Technical Data Stewards in accordance with these works.

4.1 Procedure Model

Figure 3 illustrates the overall procedure model for establishing company-wide transparency on fundamental information objects. To a large extent, the first part of the METIO procedure model is based on established approaches from systems/requirements

analysis ([Kr96], [Ba00], [So07]). The analysis process described therein was adapted to the specific requirements of our case and supplemented with corresponding roles and techniques.

The process starts with the identification of relevant information objects. This first step has to be executed in a combined top down and bottom up approach that allows for integrating essential entities both from a process and a system perspective. The most substantial entities derived from these analysis tasks are either identifying data objects without an equivalent business object on a process level, or, the other way around, business object with no analogue data object in a company’s IT systems. Moreover, business objects with multiple representations on the system level constitute further entities relevant for consolidation. For the purpose of a revelation of these misfits, a consciously separated execution of these two business tasks is recommended.

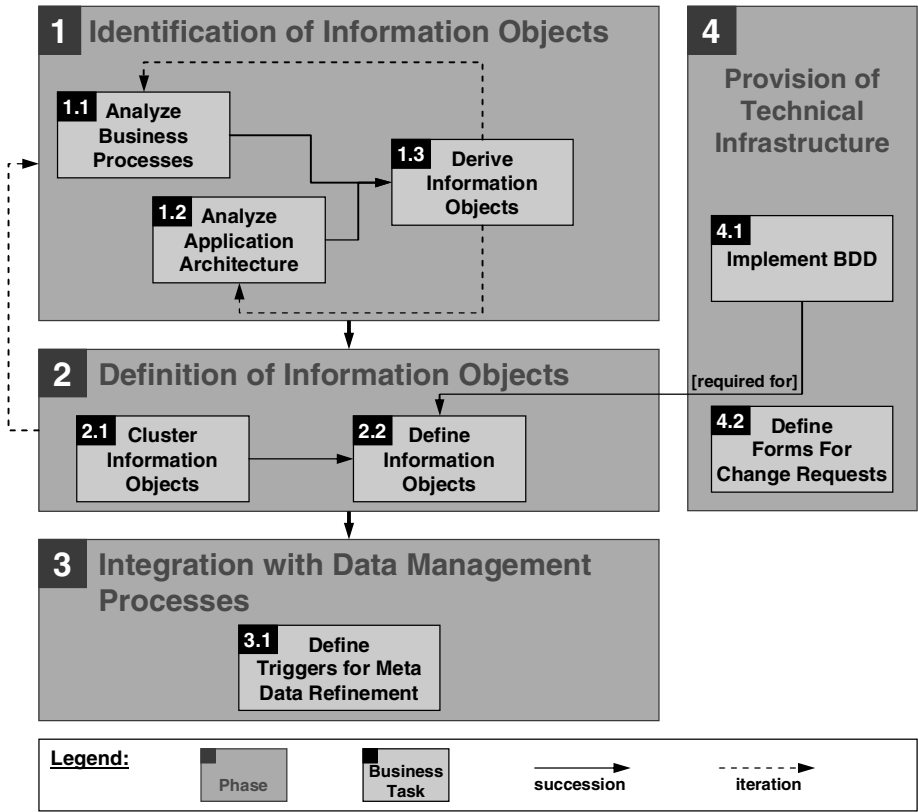


Figure 3: Procedure model of the METIO method

For the analysis on a process level, the already existing process documentation (particularly process models) needs to be worked through by the Business Data Steward (BDS) who needs to possess a sufficient understanding of the business process. If the output resulting from or being exchanged between business processes or process steps, such as business documents or goods, is mapped, these entities constitute candidates for possible

business objects. In case of insufficient process documentation (additional) interviews with Process Owners and Engineers are necessary to obtain the missing information and essential business objects.

Simultaneously, from a system perspective and complementary to the identified business objects essential data objects are to be identified by the Technical Data Steward (TDS) in collaboration with Application Owners and Data Engineers. As companies, particularly multinational enterprises, dispose of a variety of different systems and applications, real-world objects are often represented in a non-consistent way. This leads to a multitude of synonyms and homonyms inhibiting transparency and consistency. Consequently, it is necessary to consolidate the variety and obtain a redundant free mapping of the data objects on a system level. Therefore, application-specific data models as well as interviews with respective Data and Application Owners are helpful information sources.

The first phase concludes with a joint consolidation of the relevant entities resulting from the first two business tasks that needs to be executed by all Data Stewards in order to derive a consolidated and non-redundant set of information objects for definition. Hence, possible misfits between the process perspective (business objects) and the application perspective (data objects) need to be resolved. Within this business task further business or data objects might be detected, necessitating a new iteration of the before-mentioned tasks.

In order to constrain the effort for the definition of information objects, the Chief Data Steward (CDS) needs to cluster the identified entities based on their similarities. For each of these information object clusters a separate definition process variant is defined depending on the roles that are necessary to be included in the definition process. The preferable output of this first business task of the IO definition process would be a document revealing the information objects identified during the first phase, the different categories these information objects belong to, the corresponding definition process variants for the IO categories and the assignment of the roles responsible for effecting the definition (filled in by the CDS at the beginning of the definition process). Criteria for IO classification can be the usage scope of the IO (company-wide or not), whether they are used by multiple applications, their sensibility or their relevance for compliance.

This central business task for establishing transparency on information objects comprises the actual definition of the essential entities. Therefor, different definition process variants depending on the information objects cluster are possible (see Figure 4). The predefined process variants as well as the executing roles are assigned in a first step to each of the information objects. An appropriate document helps the CDS for this assignment. Each of the information objects can be specified according to the process variant it is assigned to.

When the definition of an information object is finalized, an intensive review of the result by the CDS is necessary in order to ensure completeness of the specification as well as consistency with other definitions. Finally, the Data Governance Board needs to approve and clear the definition based on the CDS's evaluation and an additional joint review. The final two activities of the business task "Define Information Objects" are

intended to aver the need for governance that emerges when information objects are to be defined on an organization-wide level. This control is needed to reconcile terms cross-functionally, with other groups in the organization who may have a different usage of a term [Ne05b].

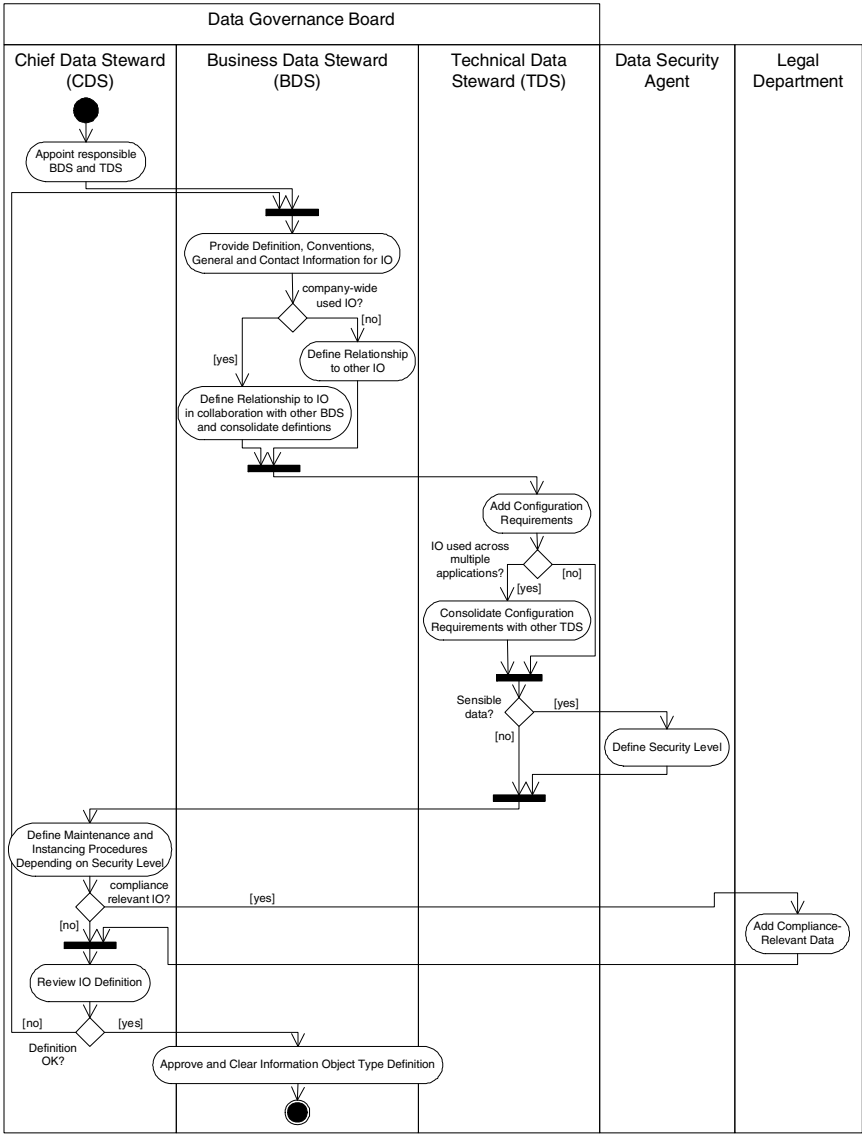


Figure 4: Activity diagram for the definition of information objects

Process step 3.1 is intended to guarantee the embedding of the metadata definition and maintenance process with the operational (meta-)data usage processes once the initial definition process is finalized. Transparency on information objects within a company cannot be realized by simply defining all relevant entities once in an unambiguous way once. In fact, it is equally important to ensure that the established transparency and consistency is surveyed continuously in order to keep a good quality of the defined information objects. This is particularly true in the dynamic environment in which companies operate nowadays: new products are launched, regulations change, mergers and acquisitions lead to new business vocabularies. And when business changes, this may lead to definitions which have been correct at one point in time but become obsolete over time. Hence, enterprises need to be able to change and adapt the definitions of relevant information objects or add new ones in the most flexible way possible [Ne05b]. This includes the possibility to make sure that conflicts are resolved and the wording of the definition is kept accurate or rapidly decide on a necessary status change of existing information objects.

For this reason we consider the process of establishing and maintaining a maximal transparency with the help of metadata as a nested and iterative process within the regular processes on a data level as illustrated in Figure 5.

This means that the triggers for transition to and from the metadata definition process (from the dark grey cycle to the nested light grey one) are required to be clearly defined. Therefore, it needs to be ensured that the preconditions for re-entering the metadata definition process in order to adapt and refine existing information object definitions, or integrate new ones, are regularly checked.

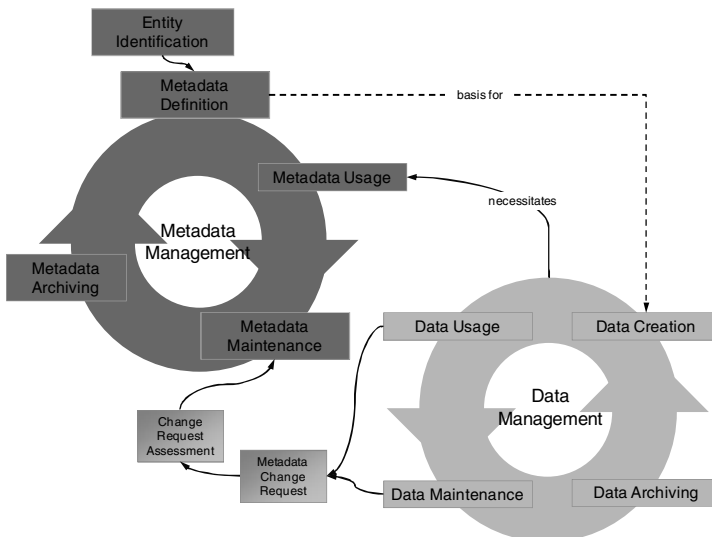


Figure 5: Metadata and data management in a conjoint lifecycle

Finally, process steps 4.1 and 4.2 are supporting activities necessary to implement the technical infrastructure, particularly the metadata repository (BDD). However, these process steps will not be detailed in this paper due to space limitations as they are not considered vital for a fundamental understanding of METIO.

4.2 How to Define an Information Object

For the BDD definitions in our prototype a number of attributes can be maintained in order to allow for a comprehensive and unambiguous specification with a maximum of semantic information. Regarding this set of attributes, the question arises why we included exactly these metadata elements (and not others as well) and whether they actually allow us a sufficient and unambiguous definition of each information object. Therefore, we conducted an in-depth literature analysis including metadata standards in adjacent research fields such as computer as well as library and information science where metadata play an important role. These potentially relevant attributes were then discussed with domain experts and complemented with further characteristics that are important from a business perspective.

PÄIVÄRINTA ET AL. identified a set of metadata elements as a result of their study of 19 contemporary public standards and specifications for document management that were considered potentially relevant [PTY02]. From the entirety of stated metadata elements (i.e. attributes) the authors extracted the ones stated most often in the standards and refined them by adding attributes from organizational needs obtained in discussion with representatives from the domain. Most of the 14 metadata elements, so-called “core elements”, derive from the Dublin Core Metadata Specification [DC08], the Australian Government Locator Service [AG02] and the ISO/IEC 11179-3 and -4 (specification and standardization of data elements and formulation of data definitions) ([ISO95], [ISO03]). The identified metadata elements are summarized in alphabetical order in Table 1. Besides, we included a short description and their equivalents from our BDD.

Element name	Description	BDD attribute
Data type	Data type of a metadata element (e.g. character string)	Data Type and Field Length
Default value	Default value of a metadata element	---
Definition	Short description of a metadata element; what is the content of the element	Definition
Example	Examples of the values assigned to a metadata element	Potential Values
Identifier	Unique identifier of a metadata element	Provided by an unambiguous name
Max. occurrence	Number of values assigned to a metadata element. The repeatability of the metadata element.	---
Name	Name of the metadata element	Name
Obligation	Obligation of a metadata element: mandatory (M),	---

	conditional (C) or optional (O)	
Producer(s)	Organization/department/team/person/role, that produces the content of a metadata element and is responsible for it	Responsible Business Data Steward
Purpose and comments	Justification; why is this metadata element needed? How is it used? Other comments or instructions.	<ul style="list-style-type: none"> ○ Rationale ○ Comment
Standard	Standard or specification, which defines the metadata element in question (name of standard and element).	---
Sub-elements	Sub-elements of a metadata element	Related Terms/Relationship
User(s)	Organization/department/team/person/role, that uses a metadata element	Where Used
Value qualifier	Name of the set of values or list of values that can be assigned to a metadata element. There can be one or more sets of values	Potential Values

Table 1: Attributes to describe metadata according to [PTY02] and their equivalent BDD attributes

As Table 1 indicates, we used the majority of the identified attributes for our BDD, however we adapted some of the elements with regard to their name and scope. The attribute “Purpose and comments” was split into two separate items and the first renamed into “Rationale” as this represents the underlying semantics more adequately. Attributes, such as “Max. occurrence” or “Default value”, were omitted due to their minor relevance for our cases.

A similar synthesis was conducted by O’NEIL for the components necessary to provide a sound definition within a glossary [Ne05a]. From this list of attributes we identified a number of further components that we could use for the BDD in addition to the ones stated above (such as “Name” and “Examples”). The attributes appended are:

- “Related”, “Narrower” and “Broader Term” were subsumed under the generalized/aggregated term “Related Terms” that – in our case – incorporates the relationship to other information objects and can be a “is-a” (corresponding to a broader-narrower-term relationship) as well as a “see also” relationship;
- “Source” was slightly changed in its naming (to “Strategic Source”) and meaning, signifying the original source of the information object rather than the source where the definition came from;
- “Approval Information” as an attribute for tracking the governance trail was renamed to “Internal Definition Process” but kept with the same semantic meaning (comprising the approval process for the definition); and
- “Distinguishing Characteristics” and “Synonyms” were directly transferred to our own BDD model with the definitions as stated in [Ne05a].

The element “Replaced by” was conceptualized broader and realized in a slightly different way. It constitutes a possible value within the BDD attribute “Status” (with Draft, Final and Retired being other possible status values). In case of a “Replaced” status of an information object a relationship “replaces/replaced by” has to be assigned to the attribute “Related Terms” in order to correctly map the replacement of one entity with another.

As those studies clearly lack a specific business and implementation focus, the results could not be transferred directly to our BDD and had to be either adapted to our specific needs (as outlined above) or supplemented by further attributes. For this purpose we integrated the information gathered from interviews and discussions with domain experts from our research project to allow an implementation that serves the requirements of our project partners. This enabled us to complement the results of the literature review with their tacit experience and knowledge of the business context.

The elements added as a result of these interviews are either relevant for implementation (such as “Security Classification” referring to the security level etc.) or provide information for the embedding in a specific business context (such as “Subject Area”, “Scope” (of application within the organization), “Coding and Descriptive Conventions”). The relevance of attributes addressing security classification and encoding descriptions is also reflected in the metadata standards comparison by BURNETT ET AL. [Bu99]. As information has to be considered within the context of the processes and applications in which they are used, the corresponding information can be maintained in the BDD under the attributes “Usage in Processes” and “Usage in Applications”. Particularly the latter is needed within the scope of application architecture planning and development when certain applications are to be replaced or deprecated.

Lastly, we added three attributes that specify how each information object is maintained (“Maintenance Procedure” and Maintenance Process Documentation”) and instantiated (“Instancing Process Documentation”) in order to help to keep the transparency and consistency on a constantly high level.

The attributes were clustered into categories based on similarities as regards content and role assignment for the definition process. Moreover, the justification for the categorization results from practical reasons as the categories were directly used for the implementation of the metadata entry masks addressing different user needs. Figure 6 summarizes the attributes that need to be defined for unambiguously describing information objects according to METIO.

Two attributes of the BDD metamodel are of particular importance. Firstly, the attribute “Distinguishing Characteristics” offers the possibility to include pertinent characteristics with specific values for each information object being defined. The attribute should not serve as a container for all existing properties but rather those characteristics that are specific to the information object being defined. This allows demarcating entities of the BDD more clearly from each other. Secondly, the exact characterization of the dependencies to other information objects is obtained by the attributes “Synonyms” (i.e. information objects with similar semantics) and “Related Terms”. The latter is used to pre-

cisely describe the relationship to associated entities in order to provide structural information. Consequently, these attributes realize the added value of our BDD in comparison to simple glossaries.

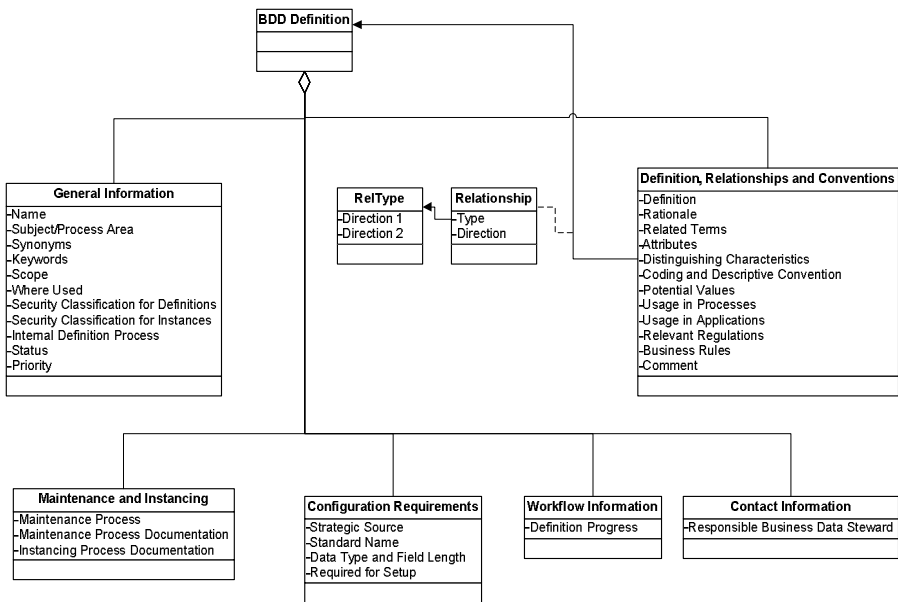


Figure 6: Attributes to be defined for comprehensive IO description in a BDD

5 Summary and Outlook

The paper at hand introduced a method that is intended to enable companies to increase the transparency and consistency among information objects on a conceptual level. Therefore, relevant entities need to be identified and then precisely defined. The (positive) consequences of unambiguously defined information objects are manifold: Firstly, they ensure a common understanding of important information objects for all entity users increasing significantly their productivity of work due to decreased search times or incorrectly stored data. Secondly, they directly increase data quality within an enterprise as all instances of used information objects are specified according to a uniform and consent definition. This, thirdly, leads to a facilitated communication with people speaking literally the same language, and helps the business make more accurate decisions [Ne05a]. And lastly, they are a prerequisite for semantic integration of heterogeneous applications that need to communicate with each other and exchange data.

Based on the action research approach introduced in the second chapter, we are currently in the third step (Action Taking) of the research process, developing the Business Data Dictionaries according to the METIO method in both real-life scenarios (see Figure 1 in chapter 2). The execution as well as the Diagnosis and the Action Planning before were

carried out in close collaboration with practitioners from the corresponding companies. After finalizing the adoption of our method within both organizations, it will be essential to evaluate the results of the implemented metadata repository and the metadata management processes, which represents an integral part of our future research.

Moreover, for further research, we consider the following issues as potential areas of interest:

- Elaboration of a metric for quantifiable evaluation of METIO,
- Utilization of metadata for data quality measurement,
- Metadata integration in other applications that are used by the ordinary employee in order to facilitate his work and improve the quality of his results,
- Extension towards an ontology-based, company-wide semantic web that allows for semantically enriched, intelligent search and real knowledge management.

6 References

- [AG02] Australian Government Locator Service: AGLS Metadata Element Set – Part1: Reference Description. Version 1.3. National Archives of Australia, Canberra, Australia, 2002.
- [Ba00] Balzert, H.: Lehrbuch der Software-Technik. Spektrum Akademischer Verlag, Heidelberg, 2000.
- [BD07] Berson, A., Dubov, L.: Master Data Management and Customer Data Integration for a Global Enterprise. McGraw-Hill, 2007.
- [BP99] Baskerville, R. L.; Pries-Heje, J.: Grounded action research: a method for understanding IT in practice. *Accounting Management And Information Technologies*, Vol. 9, pp. 1-23, 1999.
- [Bu99] Burnett, K. et al.: A Comparison of the Two Traditions of Metadata Development. *Journal of the American Society for Information Science*, Vol. 50, No. 13, pp. 1209-1217, 1999.
- [CH98] Checkland, P.; Holwell, S.: Action Research: Its Nature and Validity. *Systemic Practice and Action Research*, Vol. 11, No. 1, pp. 9-21, 1998.
- [Ch06] Chisholm, M.: Master Data versus Reference Data. *DM Review*, Vol. 16, No. 4, 2006.
- [DC08] Dublin Core Metadata Initiative: Dublin Core Metadata Element Set, Version 1.1. <http://dublincore.org/documents/dces>, accessed on: 04 February 2008.
- [DL06] Dyché, J.; Levy, E.: Customer Data Integration. John Wiley & Sons, New Jersey, 2006.
- [En99] English, L.: Improving Data Warehouse and Business Information Quality. John Wiley & Sons, New York, 1999.
- [Fr06] Friedman, T.: Gartner Study on Data Quality Shows That IT Still Bears the Burden. Gartner Group, Stamford, 2006.

- [Gu94] Gutzwiller, T.: Das CC RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen. Physica, Heidelberg, 1994.
- [He93] Heym, M.: Methoden-Engineering – Spezifikation und Integration von Entwicklungsmethoden für Informationssysteme. Universität St. Gallen, Hallstadt, 1993.
- [He04] Hevner, A. R. et al.: Design Science in Information Systems Research. Management Information Systems Quarterly, Vol. 28, No. 1, pp. 75-105, 2004.
- [ISO95] International Organization for Standardization: Information Technology – Metadata Registries (MDR). Part 4: Formulation of Data Definitions. International Standard ISO/IEC 11179-3. Geneva, Switzerland, 2003.
- [ISO03] International Organization for Standardization: Information Technology – Metadata Registries (MDR). Part 3: Registry Metamodel and Basic Attributes. International Standard ISO/IEC 11179-3. Geneva, Switzerland, 2003.
- [Kr96] Krallmann, H.: Systemanalyse im Unternehmen: Geschäftsprozessmodellierung, partizipative Vorgehensmodelle, objektorientierte Analyse. Oldenbourg, München, 1996.
- [Ma00] Marco, D.: Building and Managing the Meta Data Repository: A Full Lifecycle Guide. John Wiley & Sons, New York, 2000.
- [MS06] Marco, D.; Smith, A. M.: Metadata Management & Enterprise Architecture: Understanding Data Governance and Stewardship. DM Review, Vol. 16, No. 9-11, 2006.
- [Ne05a] O’Neil, B. K.: Business Metadata: How To Write Definitions. <http://www.tdan.com/i032fe01.htm>, accessed on: 27 January 2008.
- [Ne05b] O’Neil, B. K.: Launching a Corporate Glossary. <http://www.b-eye-network.com/view/1014>, accessed on: 27 January 2008.
- [Ni08] Nilsson, A.: Management of Technochange in an Interorganizational E-government Project. Proceedings of the 41st Hawaii International International Conference on Systems Science (HICSS-41 2008), 2008.
- [NH05] Hansen, H. R.; Neumann, G.: Wirtschaftsinformatik 1, Lucius & Lucius, Stuttgart, 2005.
- [NL06] Newman, D.; Logan, D.: Achieving Agility: How Enterprise Information Management Overcomes Information Silos. Gartner Research, Stanford, 2006.
- [OI03] Olson, J.: Data Quality – The Accuracy Dimension. Morgan Kaufmann, San Francisco, 2003.
- [OM03] Open Management Group: MDA-Guide – Version 1.0.1. 2003, <http://www.omg.org/docs/omg/03-06-01.pdf>, accessed on: 04 February 2008.
- [Pi05] Pierce, E.M.: Introduction. In: Wang, R.; Pierce, E.M.; Madnick, S. (Eds.): Information Quality. M.E. Sharpe, Armonk, NY, 2005, pp. 3-17, 2005.
- [PTY02] Päivärinta, T.; Tyrväinen, P.; Ylimäki, T.: Defining Organizational Document Metadata: A Case Beyond Standards. Proceedings of the 10th European Conference on Information Systems (ECIS), Gdansk (Poland), 2002.
- [Sc01] Scheer, A.-W.: ARIS – Modellierungsmethoden, Metamodelle, Anwendungen. Springer-Verlag, Berlin et al., 2001.
- [Sc05] Schwinn, A.: Entwicklung einer Methode zur Gestaltung von Integrationsarchitekturen für Informationssysteme, Universität St. Gallen, Difo Druck, Bamberg, 2005.

- [So07] Sommerville, I.: Software Engineering. Pearson Studium, München, 2007.
- [SZW03] Shankaranarayan, G.; Ziad, M.; Wang, R. Y.: Managing Data Quality in Dynamic Decision Environments: An Information Product Approach. *Journal of Database Management*, Vol. 14, No. 4, pp. 14-32, 2003.
- [To99] Tozer, G.: Metadata Management. Artech House Computing Library, Norwood, Massachusetts, 1999.
- [We01] Wedekind, H.: Bestandsdaten, Bewegungsdaten, Stammdaten. In: Mertens, P. et al. (Eds.): *Lexikon der Wirtschaftsinformatik*, Springer, Berlin, 2001, pp. 72.
- [We07] Wende, K.: A Model for Data Governance – Organising Accountabilities for Data Quality Management. *Proceedings of the 18th Australasian Conference on Information Systems*, Toowoomba, Australia, 2007.
- [Wh06] White, A. et al.: Mastering Master Data Management. Working Paper, Gartner Group, Stamford, 2006.
- [WH07] Wilde, T.; Hess, T.: Forschungsmethoden der Wirtschaftsinformatik. Eine empirische Untersuchung. *Wirtschaftsinformatik*, Vol. 49, No. 4, pp. 280-287, 2007.

Ein Plädoyer für die Berücksichtigung von Semantik beim Stammdaten-Alignment – Vorgehensmodell und prototypische Anwendung im Einzelhandel

Axel Winkelmann, Martin Matzner, Oliver Müller, Jörg Becker

Institut für Wirtschaftsinformatik
Westfälische Wilhelms-Universität Münster
Leonardo-Campus 3
48149 Münster

{axel.winkelmann | martin.matzner | oliver.mueller | joerg.becker}@ercis.uni-muenster.de

Abstract: Eine Automatisierung von Prozessen mithilfe einer durchgehenden informationssystematischen Unterstützung stellt hohe Anforderungen an die Qualität der zugrundeliegenden Stammdaten. In der Domäne „Handel“ adressieren Händler und Produzenten mangelnde Stammdatenqualität mit Projekten zur Vermeidung syntaktischer Datenfehler. Auch die Forschung widmet sich vornehmlich Methoden und Techniken zur Entdeckung, Vermeidung und Korrektur syntaktischer Datenfehler (z. B. fehlender Werte, Tippfehler). Semantische Belange wurden und werden bislang vernachlässigt. Unzureichende semantische Datenqualität kommt in Abweichungen zwischen relevanten Realweltentitäten, deren Attributen und Beziehungen einerseits und ihrer Repräsentation durch Daten andererseits zum Ausdruck. Semantische Datenqualität bezieht sich primär auf die gegenwärtige (fachliche) Verwendung gespeicherter Daten. Sie ist jedoch nicht auf die aktuelle Nutzung der Daten und ihren Einfluss auf bestehende Geschäftsprozesse beschränkt, sondern beinhaltet darüber hinaus auch mögliche künftige Verwendungsformen der Daten. Dieser Beitrag untersucht, inwiefern Ontologien ein geeignetes Instrument zur Identifikation und Bewertung semantischer Stammdatenfehler darstellen. Ein konzeptueller Ansatz und ein Vorgehensmodell zur Nutzung von Ontologien zur Erhöhung semantischer Stammdatenqualität werden entwickelt. Als beispielhaftes Anwendungsszenario dient das automatische Coupon-Clearing an der Kasse im Einzelhandel.

1 Motivation

Unternehmen in Industrie und Handel sind in verstärktem Maße auf korrekte und zuverlässige Stammdaten angewiesen. Im Handel werden Artikelstammdaten begleitend zu den Produkten unmittelbar von den Herstellern bereitgestellt. Die Produzenten wiederum benötigen Daten vom Einzelhandel, um z. B. eine kooperationsweite Optimierung des Güterflusses innerhalb der Lieferkette unter Berücksichtigung aktueller Kundenanforderungen durchführen zu können. Um die Datenqualität zu verbessern, haben Produzenten

und Handel begonnen, Stammdaten zu harmonisieren und zentral zu organisieren [Lo05]. Außerdem haben elektronische Marktplätze, kollaborative Führungs- und Planungskonzepte wie ECR (Efficient Consumer Response) und CPFR (Collaborative Planning, Forecasting and Replenishment) und die Einführung von zentralen Stammdatenpools die Unternehmen dazu gezwungen, die Qualität ihrer Stammdaten zu überprüfen und zu verbessern. Dennoch bleiben ein unzureichender Strukturierungsgrad der Daten und eine sehr variable und häufig unzureichende Qualität der Datenobjekte selbst bedeutende Herausforderungen im Rahmen des Stammdatenmanagements. Die Qualität der Daten ist dabei nicht als „isolierte Funktion“ eines oder mehrerer betrieblicher Informationssysteme zu verstehen, sondern ist ein inhärenter und integraler Bestandteil des gesamten Business Managements [En99].

Das Problem mangelnder Datenqualität betrifft vor allem die durch die Produzenten bereitgestellten Stammdaten. Hier mangelt es derzeit noch an standardisierten Richtlinien für Struktur und inhaltliche Beschreibung [BD07]. Beispielsweise werden identische Artikel oft inkonsistent in Datenbanken gespeichert, Schlüssel werden verschiedenartig gebildet und zahlreiche Abkürzungen verwendet. Mangelhafte Datenqualität zieht dann eine Vielzahl an Fehlern bei der Ausführung von Geschäftsprozessen nach sich und führt im Ergebnis zu einer ineffizienten Nutzung von Ressourcen in Form von Personen, Geld, Material und Maschinen [En99]. Im Handel ergibt sich als eine Konsequenz, dass bislang die Einführung von flexiblen Datenstrukturen, die über eine generische Klassifikation von Waren hinausgehen, nicht möglich ist. Aufgrund dieses Mangels an Struktur und Qualität, können Artikelstammdaten selten unmittelbar in hoch automatisierten Geschäftsprozessen verwendet werden. Vielmehr ist das Mitwirken von Menschen zum manuellen Interpretieren, Aktualisieren und Restrukturieren der Daten erforderlich [Ca04].

Als wissenschaftliche Disziplin befasst sich die Semiotik mit der zeichenbasierten Repräsentation von Realweltentitäten. Mit Rückgriff auf die semiotischen Teildisziplinen schlagen PRICE and SHANKS [PS05] syntaktische, semantische und pragmatische Kriterien für die Evaluation von Datenqualität im Allgemeinen vor. Andere Forscher zeigen Methoden und Techniken zur Beseitigung so entdeckter syntaktischer und pragmatischer Qualitätsdefizite auf [Re96]. Nur wenige Forscher diskutieren jedoch konkrete Ansätze oder praktische Richtlinien für semantische Aspekte der Stammdatenqualität [CFP03]. Ferner betrachten bisherige Untersuchungen Datenqualität vornehmlich auf der aktuellen Nutzung der Daten. Datenqualität ist jedoch nicht auf die aktuelle Verwendung von Daten und deren aktuellen Einfluss auf Geschäftsprozesse beschränkt, sondern beinhaltet auch die zukünftige Nutzung der Daten [En99]. Schließlich sind die Stammdaten eines Unternehmens die Grundlage für künftige Geschäftsideen und insbesondere Dienstleistungen, die bei der Konzeption derzeitiger Informationssysteme noch keine Berücksichtigung finden konnten.

In diesem Artikel diskutieren wir den Beitrag von Ontologien zur Identifikation und Bewertung von Stammdatenqualitätsproblemen auf semantischer Ebene. Wir erläutern unseren Ansatz mit Rückgriff auf das Szenario des automatischen Coupon-Clearings in Einzelhandelsunternehmen. Bei manuellen, wenig automatisierten Coupon-Clearing-Prozessen an der Kasse sind Einzelhändler nicht auf die Nutzung von Artikelstammdaten ange-

wiesen. Das Kassenspersonal prüft Coupons durch Inaugenscheinnahme manuell und entnimmt die durch die Promotion eingeräumten Kaufkonditionen einem Aufdruck auf dem Coupon. Die Einführung von Self-Checkout-Systemen und organisierte Coupon-Betrügereien zwingen den Handel jedoch, Coupon-Clearing-Prozesse zu automatisieren. Es entstehen dadurch neue Anforderungen an die Qualität der Artikelstammdaten [Wi06].

Dieser Artikel ist wie folgt strukturiert: In Kapitel 2 diskutieren wir den Stand der Forschungsbemühungen zu Stammdatenqualität und erläutern die verschiedenen adressierten Problemdimensionen. In Kapitel 3 beschreiben wir einen Ansatz zur Vermeidung semantischer Stammdatenqualitätsprobleme bei der Verwaltung von Artikelstammdaten durch die Nutzung von Ontologien. In Kapitel 4 erläutern wir den Beitrag unserer Arbeit für Theorie und Praxis, illustrieren Beschränkungen des Konzepts und geben einen Ausblick auf künftige Forschungsaktivitäten in diesem Bereich.

2 Status-quo der Forschung zur Qualität von Stammdaten

2.1 Stammdaten

Stammdaten umfassen diejenigen Informationsobjekte, die benötigt werden, um ein unternehmensweites „Berichtssystem“ für grundlegende betriebliche Funktionen zu schaffen und aufrecht zu halten, um betriebliche Transaktionen zu erfassen und die Ergebnisse dieser betrieblichen Funktionen zu messen [BD07] [Gr05]. Im Handel werden verschiedenartige Stammdaten benötigt und verwaltet. Händler benötigen Stammdaten einerseits über Kunden und Lieferanten wie etwa Name, Adresse, Bankverbindungen etc. und andererseits über die verwalteten Artikel. Artikelstammdaten setzen sich aus Grunddaten (z. B. Artikelnummer, UPC, EAN), Listungsdaten (Listungszeitraum, Zuordnung zu Sortimenten), Bezugsdaten (z. B. lieferantenabhängige Verfügbarkeiten), Logistik-, Absatz- und Kassendaten [BUV01] zusammen.

Automatisierte Kassenprozesse greifen auf Artikelstammdaten und Promotionsstammdaten zurück. Artikelstammdaten beinhalten Informationen über Produktname, Größe, Preis etc. Zur Identifikation von Artikelstammsätzen wird auf eindeutige Bezeichner wie die UPC (Unique Product Code (USA)), die EAN (European-/International Article Number) oder die JAN (Japanese Article Number) zurückgegriffen. Promotionsstammdaten beinhalten neben relevanten Informationen für Promotionen wie dem Wert des Promotionsangebotes, Gültigkeitsbeschränkungen u. ä. auch eine Referenz zum beworbenen Produkt. Zum Beispiel benötigt der Händler zur automatisierten Abwicklung einer BOGOF-Promotion (buy one, get one free) Promotionsstammdaten über den Wert der Promotion („ein Artikel umsonst“), die Laufzeit der Promotion („bis zum 31. Juli“), und das beworbene Produkt beziehungsweise den Produktcode des Produkts (z. B. „Coca-Cola 0,5 Liter“, EAN: 4023786889760, EAN: 4046758456739, EAN:...).

2.2 Stammdatenqualität

Zahlreiche Forschungsaktivitäten ergründeten Kriterien, in denen Daten- und Informationsqualität zum Ausdruck kommt.¹ Mehrere Autoren bilanzierten in den vergangenen Jahren eine tendenziell unzureichende Stammdatenqualität in Industrie und Handel. KUIPERS berichtet zum Beispiel, Nestlé vermute, dass zu über 50% der gelisteten Produkte inkorrekte oder redundante Stammdateneinträge existierten [Ku04]. Im Rahmen einer 2005 in den USA durchgeführten Studie wurde darüber hinaus erhoben, dass 30% der Data-Warehouse-Betreiber mit massiven, durch unzureichende Datenqualität bedingten, Problemen kämpften. Für eine generellere Einschätzung zum Thema „Datenqualität“ sei auf die Arbeiten von [Ag05] sowie [BW08] verwiesen. Eine aktuelle Untersuchung zu den Defiziten bei der Datenqualität von Verbundgruppen im Handel findet sich bei [BGW08].

Eine Quintessenz der Forschungsaktivitäten der vergangenen zwei Jahrzehnte ist, dass es sich bei Datenqualität um ein „multidimensionales Konzept“ handelt; vgl. [PLW02] und die dort genannten Quellen. Es lassen sich unterscheiden: Eine (a) *subjektiv* empfundene Datenqualität, die in der Einschätzung der Datennutzer zum Ausdruck kommt, und eine (b) *objektive* Bewertung der betrachteten Datensätze anhand akzeptierter Qualitätskriterien. Andere Autoren beschreiben diese Phänomene mit den Ausdrücken *pragmatische* (a) und *inhärente* (b) Datenqualität [En99]. Qualität auf pragmatischer Ebene bezeichnet Korrektheit im Sinne einer korrekten Repräsentation (a.I) „richtiger“ Fakten [En99]. Das Hauptaugenmerk liegt hier auf der Bereitstellung und schließlichen Nutzung der Datenobjekte durch den Anwender. Qualität im Sinne inhärenter Datenqualität ist beschränkt auf die Korrektheit (b.II) der gespeicherten Fakten und eine klare Definition (b.III) und damit Bedeutung der Daten. Die korrespondierenden Subkategorien der semiotischen Lehre sind *Syntax* und *Semantik* [PS05]. Für alle drei Komponenten (Definition, Korrektheit der Fakten und Repräsentation) können voneinander losgelöst Maßnahmen zur Erhöhung der Datenqualität durchgeführt werden [En99]. Eine separate Messung der Qualitätskomponenten wird durch Interrelationen hingegen verhindert [LSK02]. Der folgende Abschnitt adressiert die semantischen Aspekte von Datenqualität und fokussiert dabei die Datendefinitionskomponente (b.III) semantischer Stammdatenqualität.

2.3 Semantische Stammdatenqualität

Semantische Stammdatenqualität ist ein Ausdruck für das Maß an Übereinstimmung zwischen den Daten, die in einer Datenbank abgelegt sind, und der Summe an Eigenschaften externer Realweltphänomene, die sie repräsentieren sollen und die für den mit der Datenspeicherung verbundenen Zweck relevant sind [PS05]. Eine etablierte Methode zur Bewertung *semantischer Datenqualität* ist das Random Sampling dieser Abbildungen. Datenqualität auf semantischer Ebene in diesem Sinne ist ein Ausdruck der Über-

¹ Die Begriffe Daten- und Informationsqualität werden hier synonym verwendet; vgl. dazu [HLW99] [KPM04]).

einstimmung einer epistemologischen Sicht auf die Datenobjekte mit der korrespondierenden ontologischen Beschreibung der externen Phänomene [HKL95].

Im Gegensatz zu diesem umfassenden Verständnis, beschränken viele Autoren Aspekte semantischer Qualität auf die Vermeidung von Inkonsistenzen, die durch redundante Speicherung von Informationselementen auftreten (z. B. auch CODD in seinem wegweisenden Beitrag [Co70]), vernachlässigen jedoch (a) das korrekte Mapping von Daten zu den Realweltentitäten während des gesamten Lebenszyklus von Realweltentitäten und Daten, und (b) die Eignung der Daten zur Verwendung in künftigen neuen fachlichen Kontexten.

(a) Korrektes Mapping: Das Postulat des korrekten Mappings korrespondiert mit dem Konzept des „Information as a Product“ [Pi04] [Wa98] [WYP98]. Datenqualität kann hier einerseits durch die Anwendung von Methoden sichergestellt werden, die die Konformität der Daten zu einem initial spezifizierten Anforderungskatalog und weiteren Integritätsregeln garantieren. Alternativ können externe Phänomene z. B. mithilfe von Ontologien [PS05] fachlich beschrieben und die Entsprechung mit den gespeicherten Daten geprüft werden. Ebenso wie physikalische Produkte folgen Daten einem Lebenszyklus. Dabei ist die Veränderung der Daten im Zeitablauf eine Funktion der Informationssysteme, die die Unternehmensdaten erstellen und ändern [HLW99]. Der Datenlebenszyklus beschreibt Erstellung, Veränderung und Löschung der Datenobjekte.

(b) Nutzung im neuen Kontext: Nicht nur die Stammdaten, auch die abgebildeten Realweltentitäten, die Geschäftsregeln, die diese miteinander verbinden, und der umgebende fachliche Anwendungskontext sind dynamisch. Folglich kann die spätere operative Nutzung der Daten, bedingt durch nicht berücksichtigte, unvorhergesehene oder veränderte fachliche Anforderungen, sich grundlegend von derjenigen unterscheiden, die während der Systementwicklung spezifiziert wurde [PS05]. Diese Aspekte sollen unter dem Begriff *kontextuelle Datenanforderungen* subsumiert werden. Beispiele im Handel sind umfassende Dokumentations- und Berichtspflichten, die z. B. die Verwendung von Inhaltsstoffen (z. B. genetisch veränderte Inhaltsstoffe), Gefahrgütern (z. B. REACH, ROHS oder WEEE) oder artikelspezifischen Informationen zur Erstellung finanzwirtschaftlicher Berichte betreffen. Gerade hier sehen sich Unternehmen mit zahlreichen neuen Gesetzen und Verordnungen (z. B. Sarbanes-Oxley, Basel II) konfrontiert, die sie zur Bereitstellung, Nutzung und zum Reporting aktueller, verifizierbarer und relevanter Daten über die allgemeine betriebswirtschaftliche Leistungsfähigkeit und signifikante Ereignisse zwingen [BD07].

Zum Verständnis von kontextabhängigen Aspekten der Stammdatenqualität sind die Metadaten der Datenobjekte zu berücksichtigen. Die Metadaten beinhalten Definitionen und Dokumentationen, die einerseits die Domäne der fachlichen Nutzung in der Realwelt und andererseits das Datenmodell beschreiben [PS05]. Sie adressieren somit gezielt die zuvor identifizierte Lücke zwischen Realweltbeschreibung und Datenbankinhalt.

2.4 Forschungsmethodischer Ansatz

Unsere Arbeit untersucht das Konstrukt „Ontologiebasiertes semantisches Stammdatenmanagement“. Als qualitative Forschungsmethode wurde „Design Science“ zur Führung des Forschungsprozesses gewählt. Nach HEVNER U. A. versucht Design-Science-Forschung, IT-Artefakte zu erstellen und zu evaluieren, um identifizierte organisatorische Probleme zu lösen [HMP04]. Für die Ermittlung und Beschreibung dieser Probleme gilt es, im Rahmen eines Design-orientierten Forschungsprozesses insbesondere, Bedeutung und Relevanz der adressierten Problembereiche für die untersuchte Domäne zu belegen. Dazu kann auf eine Vielzahl von Forschungsmethoden wie Literaturanalyse, fallstudienbasierte Untersuchungen und Experteninterviews [Ja91, Mi97, Sc91] oder eine pluralistische Kombination dieser Ansätze zurückgegriffen werden [Pe07]. Die in diesem Artikel dargelegten Erkenntnisse basieren auf praktischen Erfahrungen aus zahlreichen Experteninterviews und sechs Stammdatenmanagementprojekten, die in den vergangenen Jahren am Institut begleitet wurden. Der Artikel beschreibt die Weiterentwicklung des in Zusammenarbeit mit einem Clearingunternehmen entwickelten ontologischen Stammdatenmanagementansatzes, der auf den Erfahrungen aus den genannten Projekten (vgl. Tabelle 1) aufsetzt (vgl. auch [BJ07a, BJ07b, BJ07c, Wi06]).

Fallstudie	Unternehmen	Durchgeführte Projekte
Handelsunternehmen	Großhändler mit 400 Mitarbeitern, 50 Großmärkten und annähernd 30.000 Artikeln.	Prozessanalyse, Stammdatenreorganisation sowie Auswahl und Analyse eines neuen ERP-Systems.
Handelsunternehmen	Juwelier- und Uhrmacherhandelskette mit 200 Geschäften und 2.000 Mitarbeitern in ganz Deutschland.	Reorganisation des Controlling-Systems mit Vereinheitlichung des Stammdaten- und Kennzahlenkonzepts. Einführung eines Data-Warehouse.
Gebäudeserviceunternehmen	Komplettserviceanbieter für Immobilienverwaltung. Management von mehr als 35.000 Objekten, 6.800 Mitarbeiter.	Einführung eines neuen Stammdatenkonzepts.
Industrie	Kleines metallverarbeitendes Unternehmen mit 125 Mitarbeitern. Entwicklung von Produkten mit einem hohen Grad an manueller Fertigung.	Prozessautomatisierung und Stammdatenreorganisation. Einführung eines Produktionsplanungssystems.
Betreiber eines europäischen Artikelstammdatenpools	Hersteller von Lösungen für das Management von strukturierten und unstrukturierten Daten, multimedialen Inhalten und Transaktionsdaten von weit verteilten Quellen.	Analyse des existierenden Stammdatenpools und seines syntaktischen Qualitätsmaßes.
Deutsches Coupon-Clearing Unternehmen	Zehn Mitarbeiter, vereint ca. den halben deutschen und große Teile des europäischen Marktes für automatische Coupon-Prozess-Lösungen im Handel auf sich.	Entwicklung einer in-store Coupon-Prozess-Lösung. Entwicklung eines zentralisierten Stammdatenpools für Promotionsstammdaten.

Tabelle 1: Berücksichtigte Fallstudien und Projekte zum Stammdatenmanagement

3 Ontologie-basiertes, semantisches Stammdatenmanagement

Dieser Abschnitt beschreibt einen Ansatz zum Management semantischer Stammdatenqualität. Wie bereits dargelegt, kann nach PRICE UND SHANKS [PS05] syntaktische Datenqualität durch Integritätsüberprüfung, semantische Datenqualität durch Random Sampling und pragmatische Datenqualität – da sie entschieden von der Wahrnehmung des Benutzers abhängt – nur durch empirische Methoden, wie z. B. Befragungen oder Interviews, überprüft werden. Im Folgenden stellen wir einen systematischen Ansatz zur Messung der semantischen Datenqualität vor, der über die zufällige Stichprobengenerierung des Random Sampling hinaus geht.

Die Grundidee ist, ein Informationssystem als Repräsentation eines Realweltssystems zu verstehen [We97]. Ein Zustand des Realweltssystems zu einer bestimmten Zeit wird von den Daten repräsentiert, die im Informationssystem gespeichert sind [WW96]. Da Informationssysteme in der Regel nur einen unvollständigen und vereinfachten Ausschnitt der realen Welt abbilden, stellen sich zwei Fragen:

- Frage 1: Welche Realweltphänomene sollen im Informationssystem repräsentiert werden (d. h. welche Entitäten sind relevant)?
- Frage 2: Wie sollten relevante Entitäten im Informationssystem repräsentiert werden (d. h. welche Attribute und Beziehungen sind relevant)?

Die Beantwortung dieser Fragen ist eine notwendige (jedoch nicht hinreichende) Bedingung zur Beurteilung der semantischen Datenqualität eines gegebenen Informationssystems. Um dieses Problem anzugehen, schlagen wir vor, zunächst ein konzeptionelles Modell zu erstellen, das die Semantik der realen Welt, die wir repräsentieren wollen, dokumentiert. In darauf folgenden Schritten wird dieses konzeptionelle Modell mit dem physischen Datenmodell des Informationssystems und dessen Datenbestand verglichen.

Ein möglicher Ansatz zur Abbildung der Semantik einer Domäne ist die Erstellung einer Ontologie. Der Begriff Ontologie entstammt der Philosophie und bezieht sich auf die Studie des Seins. Die zentrale Fragestellung dieser Disziplin lautet: „Was existiert?“ oder um präziser zu sein [So00] „Welche Kategorien von Dingen existieren?“. Eine Domänenontologie spezifiziert folglich die fundamentalen Typen von Entitäten innerhalb einer wohl abgegrenzten Domäne. Im Kontext von Informationssystemen wird eine Ontologie häufig als ein abstrakter, vereinfachter Blick auf die Welt, die wir für einen bestimmten Zweck repräsentieren möchten, definiert [Gr93]. Hier besteht eine Ontologie aus einer Menge von Konstrukten (representational primitives), mit denen eine Domäne modelliert werden kann (vgl. im Folgenden [Gr08]). Bei diesen Konstrukten handelt es sich typischerweise um Konzepte (häufig auch Klassen genannt), Attribute dieser Konzepte (häufig auch Eigenschaften genannt) und Beziehungen zwischen Konzepten. Die Definitionen der grundlegenden Konstrukte umfasst Informationen über ihre Bedeutung und Bedingungen ihrer logisch konsistenten Anwendung. Im Kontext von Datenbanksystemen können Ontologien als ein Detaillierungsgrad, der von Datenmodellen und Implementierungsfragen abstrahiert, betrachtet werden.

Liegt eine Ontologie für eine relevante Domäne der Realwelt vor, kann dieses konzeptionelle Modell sowohl mit dem physischen Datenmodell des Informationssystems als auch mit dem genutzten aktuellen Datenbestand verglichen werden. Abweichungen, die bei dieser Analyse offengelegt werden, können als mögliche Indikatoren für mangelnde Datenqualität interpretiert werden. Abbildung 1 detailliert den geschilderten Ansatz, der in die vier Phasen Analyse, Modellierung, Mapping und Anwendung unterteilt werden kann. Jede Phase beinhaltet Aufgaben, die durch Artefakte verbunden sind. Die einzelnen Phasen sind Gegenstand der folgenden Unterabschnitte. Die abgebildeten Modellelemente und deren Beziehungen werden dort explizit adressiert.

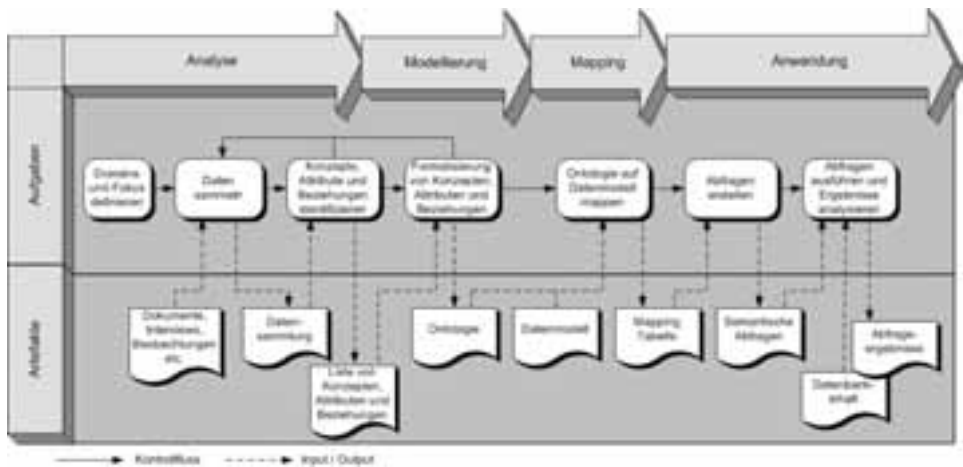


Abbildung 1: Vorgehensmodell für ein semantisches Datenqualitätsmanagement

3.1 Analyse

Der erste Schritt beinhaltet die Definition sowie klare Abgrenzung der zu behandelnden Domäne. Beispielhafte Schwerpunkte im Handel sind u. a. IT-unterstützte Geschäftsstrategien (z. B. E-Commerce), Geschäftsprozesse, die auf hohe Datenqualität angewiesen sind, (z. B. CPFR, ECR, RFID, Promotionen) oder gesetzliche Richtlinien, die Voraussetzungen für das Datenmanagement definieren (z. B. SOX, Basel II).

Im Folgenden werden wir das sogenannte Couponing als ein durchgehendes Beispiel verwenden. Coupons sind ein Marketinginstrument in Form eines gedruckten oder elektronischen Gutscheins, der dem Besitzer entweder direkt oder indirekt Rabatt zugesteht, wenn bestimmte Einlösebedingungen erfüllt sind [Wi06]. In der Vergangenheit wurden Coupons manuell am Point-of-Sale (POS) verrechnet. Heute wird automatisches Coupon-Clearing immer wichtiger, nicht zuletzt, da es viele der Nachteile des manuellen oder halbautomatischen Clearings vermeidet (insb. Zeitaufwand, Fehleranfälligkeit und Betrugsrisiko) und eine automatische Verarbeitung am POS ohne manuelle Eingriffe ermöglicht (z. B. an Self-Checkout-Automat). Das gleiche gilt für das Ausstellen von Coupons an der Kasse. Heute ist es möglich, Coupons gemeinsam mit dem Kassenschein zu drucken, wenn bestimmte Bedingungen erfüllt sind. Kauft ein Kunde zum Beispiel

einen Artikel eines bestimmten Herstellers, erhält er einen Coupon, der ihm beim nächsten Einkauf Rabatt auf andere Artikel (z. B. Neueinführungen) des gleichen Herstellers gewährt. Automatisches Couponing hängt stark von einer effektiven und effizienten IT-Unterstützung ab. Insbesondere die Qualität der benötigten Stammdaten ist eine notwendige Voraussetzung für ein automatisches Abwickeln von Coupon-Aktionen. Das Zusammenspiel von Coupon und Artikelstammdaten illustriert Abbildung 2.

Es ist ratsam, die relevante Domäne nicht nur zu identifizieren, sondern auch sorgfältig abzugrenzen. Eine bekannte Technik ist die Verwendung von so genannten motivierenden Szenarien und Kompetenzfragen ([GF95], [Us96], [UG96]). Die motivierenden Szenarien haben die Form von Geschichten, Beispielen oder Problembeschreibungen, enthalten aber auch mögliche Lösungsansätze. Diese Lösungsansätze liefern oft eine erste Idee für die beabsichtigte Semantik der Ontologie. Aus den motivierenden Szenarien kann eine Menge von Fragen, die so genannten Kompetenzfragen, abgeleitet werden. Die Kompetenzfragen decken Anforderungen bezüglich der zu entwickelnden Ontologie auf. Ihr primärer Zweck ist jedoch nicht die Generierung von ontologischen Verpflichtungen; sie sind vielmehr dazu bestimmt als eine Art Nagelprobe für die fortlaufende Evaluation der Ontologie zu dienen [NM08], z. B.: Enthält die Ontologie genug Informationen, um alle Kompetenzfragen zu beantworten? Ist das Detaillierungslevel passend, um die Kompetenzfragen effektiv und effizient zu beantworten?

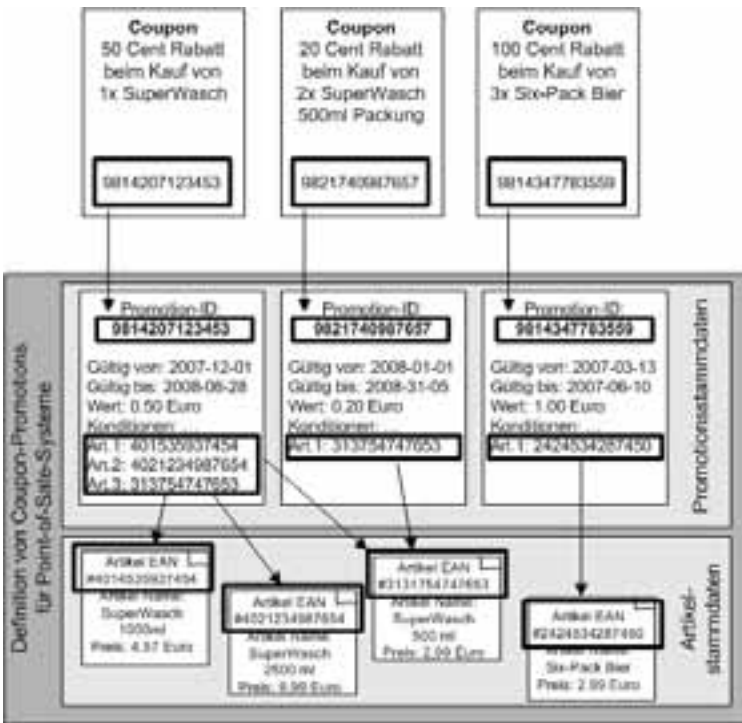


Abbildung 2: Zusammenspiel von Coupon, Barcode und Artikelstammdaten im automatischen Coupon-Clearing

Im Folgenden wird ein exemplarisches motivierendes Szenario für unser begleitendes Beispiel vorgestellt: Ein Händler will ein neues Fleckensalz bewerben. Er umwirbt Kunden, die Artikel mit fleckenverursachenden Inhaltsstoffen, wie z. B. Tomaten, kaufen. Es soll automatisch ein Coupon für eine kostenlose Packung Fleckensalz an der Kasse ausgedruckt werden, wenn ein Kunde ein Produkt kauft, das Tomaten enthält, bspw. Ketchup oder Pastasauce. Die Aktion ist auf zwei Wochen begrenzt.

Beispielhafte Kompetenzfragen:

- Welche Artikel sind aus Tomaten gemacht?
- Welche EANs (Europäische Artikelnummer) gehören zu Artikeln mit Tomaten?
- Welche EANs gehören zum Fleckensalz (oft hat ein Artikel mehr als eine EAN – abhängig vom Produktionsort etc.)?
- Läuft die Werbeaktion noch?

Ist die Domäne definiert und abgegrenzt, so folgt die Datensammlung. Eine Vielzahl von qualitativen empirischen Methoden kann dafür benutzt werden. Wichtige Datenquellen, die als Input für diese Aufgabe dienen können, sind Dokumentanalysen (z. B. von Gesetzen, Richtlinien, Geschäftsdokumenten, Geschäftsberichten, Mitschriften von Meetings), strukturierte, offene oder fokussierte Interviews (z. B. mit Domänenexperten, Mitarbeitern) und direkte oder teilnehmende Beobachtungen (z. B. von Mitarbeitern). Der Output dieser Aktivität ist eine konsolidierte Datensammlung, z. B. in Form eines zentralen Berichtes, der als Input für alle weiteren Phasen der Analyse dient. Auf der Basis dieser Datensammlung können zentrale Konzepte, Attribute und Beziehungen abgeleitet und detailliert werden. Dabei ist es hilfreich, zunächst durch Brainstorming zentrale Begriffe der Domäne zu identifizieren und sich nicht zu sehr über Klassifikationen und Überschneidungen Gedanken zu machen [Us96]. In nachfolgenden Schritten werden die gesammelten Begriffe als Konzepte, Attribute oder Beziehungen klassifiziert und ausdifferenziert, um Überschneidungen zu reduzieren. Während dieser Aufgabe kann es notwendig sein, zu vorgelagerten Schritten zurückzugehen, wenn Informationen über wichtige Aspekte fehlen oder wenn man identifizierte Konzepte, Attribute und Beziehungen evaluieren möchte. Der finale Output der Analysephase ist eine semi-strukturierte Liste von zentralen Konzepten, Attributen und Beziehungen der analysierten Domäne.

3.2 Modellierung

Der Zweck dieser Phase ist die Formalisierung der Ergebnisse der vorangegangenen qualitativen Analyse durch die Konstruktion der eigentlichen Domänenontologie. Dies bezieht typischerweise die Kodierung der identifizierten Konzepte, Attribute und Beziehungen mittels einer formalen Sprache, z. B. OWL (Web Ontology Language), RDF (Resource Description Framework) oder UML (Unified Modelling Language) und eines Modellierungswerkzeugs (z. B. Protégé) ein.

Ein sinnvoller Ansatz ist es, die Liste der identifizierten Konzepte zunächst in einer hierarchischen Taxonomie anzuordnen. Danach wird die interne Struktur der Konzepte durch die Definition beschreibender Attribute (und ihrer Wertebereiche) spezifiziert. Der

nächste Schritt ist die Definition von Beziehungen (inklusive Kardinalitäten) unter den Konzepten. Der letzte (optionale) Schritt umfasst die Erstellung von Instanzen der Konzepte, das Ausfüllen von Attributwerten sowie die Definition konkreter Beziehungen zu anderen Instanzen.

Abbildung 3 zeigt die Ontologie für unser begleitendes Beispiels in OWL-Notation. Die zentralen Konzepte sind Coupons, Promotionen, Konditionen, Rabatte (Ermäßigungen und Naturalrabatte), Artikel und Inhaltsstoffe. Zusätzlich sind Beziehungen zwischen Konzepten enthalten, z. B. Coupons gehören zu einer Promotion, Artikel beinhalten Inhaltsstoffe und Artikel sind zu EANs zugeordnet. Die Ontologie enthält auch Instanzen von Konzepten. Zum Beispiel beinhaltet der Artikel „Mama Mia Pasta Sauce“ den Inhaltsstoff „Tomate“ und die EAN „300819814711“. Analog hat der Coupon „Coupon_4711“ die Kondition „Regel_Tomate-Fleckensalz“, welche die Prämisse Inhaltsstoff = „Tomate“ und die Konklusion „GetOneFree“ des Artikels „FleckWeg_Fleckensalz“ hat.

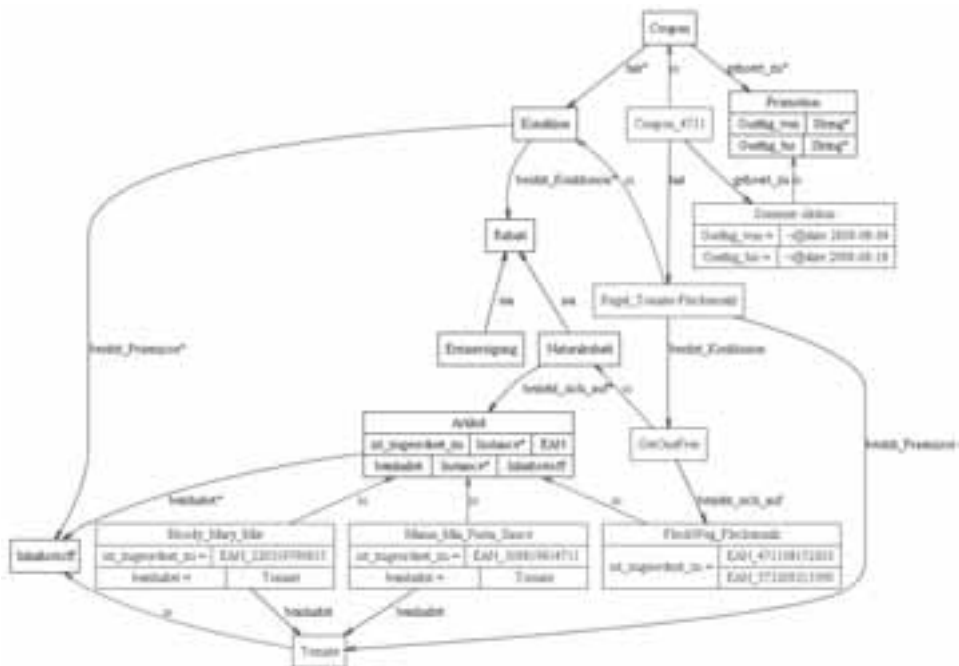


Abbildung 3: Exemplarische Ontologie (inkl. Instanzen) für Couponig-Aktionen im Einzelhandel

3.3 Mapping

Um die semantische Datenqualität eines bestimmten Informationssystems tatsächlich zu analysieren, muss ein Mapping zwischen der Domänenontologie und der zu betrachtenden Datenbank durchgeführt werden (Abbildung 4). Dieses Mapping verbindet die Semantik der Ontologie mit dem physischen Datenmodell und dem aktuellen Datenstand

der Datenbank. Für jedes Konzept (inkl. der Attribute) und alle Beziehungen innerhalb der Ontologie muss ein Gegenstück im vorliegenden Datenmodell identifiziert werden. NECIB und FREYTAG [NF05] haben in diesem Zusammenhang drei Typen von Mappings identifiziert: Zunächst Mappings zwischen Ontologiekonzepten (inkl. Attributen) und Datenbankrelationen (d. h. Tabellen und Views). Das Ontologiekonzept „Artikel“ könnte z. B. auf die Datenbankrelation „Item“ abgebildet werden. Desweiteren kann ein Mapping von Ontologiebeziehungen auf Datenbankrelationen erfolgen. Zum Beispiel könnte die Ontologiebeziehung „beinhaltet“ auf die Datenbankrelation „Rezept“ abgebildet werden, die Inhaltsstoffe von „items“ enthält. Dieser Link könnte eine Relation (im Falle einer many-to-many-Kardinalität) oder auch nur eine Foreign-Key-Bedingung (im Fall einer one-to-many-Kardinalität) sein. Abschließend können Mappings zwischen Ontologiekonzepten (inkl. Attributen) oder deren Instanzen und Datenbankattributwerten definiert werden. Zum Beispiel könnte die Ontologieinstanz „Tomate“ zum Wert „true“ des Attributes „tomatenhaltig“ in der Datenbankrelation „items“ gemappt werden.

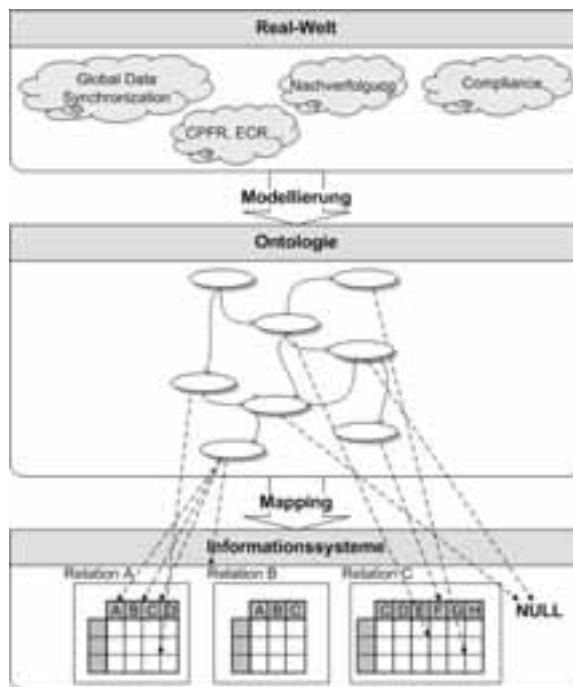


Abbildung 4: Präsentierter Ansatz zum semantischen Datenqualitätsmanagement

In einigen Fällen wird es nicht möglich sein, eine Entsprechung für ein Ontologiekonstrukt im Datenbankschema sofort zu finden. Dies kann der Fall sein, wenn ein Phänomen der realen Welt nicht im Informationssystem repräsentiert ist. Zum Beispiel besitzen Standard-ERP-Systeme nicht alle Relationen und Attribute, die im Handel benötigt werden (z. B. Promotionen oder Coupons). In diesem Fall liegt ein Defekt des Datenmodells (und nicht des Datenbestandes) vor, welcher nur durch Modifizierung oder Erweiterung des Datenmodells behoben werden kann [Sh99]. Zur schließlichen Ausführung existieren Konzepte zur Erweiterung der SQL-Funktionalität [Prud06].

Das Ergebnis der Mapping-Phase ist eine Mapping-Tabelle, die auf der einen Seite Ontologiekonstrukte enthält und andererseits Relationen und Attributwerte des Datenbankschemas ausweist.

3.4 Anwendung

Die letzte Phase hat die resultierende Qualitätsbewertung zum Gegenstand, die den Entwurf von semantischen Abfragen, die Ausführung dieser Abfragen und die Analyse der Abfrageergebnisse beinhaltet. Aufgrund der vorangegangenen Konstruktion einer Domäneontologie und des Mappings zwischen Ontologiekonstrukten und dem physischen Datenmodell, können Analysten Abfragen auf einem semantischen Level entwerfen ohne sich dabei mit Details des physischen Datenbankmodells zu beschäftigen. Diese semantischen Abfragen werden dann basierend auf der Mapping-Tabelle in SQL-Abfragen übersetzt. Nach Ausführung der Abfragen kann beantwortet werden, ob alle Attribute und Beziehungen, die in einem spezifischen Kontext, wie z. B. dem Couponing, benötigt werden, gefüllt sind. Im Gegensatz zur Beantwortung solcher Fragestellungen durch reine SQL-Abfragen ohne übergeordnete Semantik erlaubt ein ontologiebasierter Ansatz zudem die Überwindung von typischen Heterogenitätsproblemen, z. B. inkonsistente Benennung von Relationen, Attributen und Attributwerten. Abbildung 5 zeigt einen Screenshot eines ersten Prototyps, der den präsentierten Ansatz zur semantischen Bewertung von Stammdatenqualität im Bereich Couponing mit einbezieht. Die grünen Balken auf der rechten Seite des Screenshots repräsentieren den Grad der Übereinstimmung zwischen der Ontologie und dem physischen Datenmodell und Datenbestand. Das Qualitätsmaß in diesem ersten Prototyp ist relativ einfach gehalten und baut auf einem linearen Modell auf, das Inkonsistenzen und falsche Bezeichnungen berücksichtigt.



Abbildung 5: Beispielhafte Anwendung

4 Zusammenfassung und weiterer Forschungsbedarf

Fragestellungen bezüglich der syntaktischen Aspekte von Datenqualität sind in der Forschung weitgehend durchdrungen (auch wenn sie in der Praxis noch nicht zufriedenstellend gelöst sind). Praktische Ansätze bezüglich der semantischen Aspekte von Datenqualität existieren hingegen kaum.

Der präsentierte Ansatz soll zur Diskussion über Möglichkeiten der Verwendung von Ontologien für das semantische Stammdatenmanagement beitragen. Der Ansatz adressiert eine Reihe von Fragen auf dem Gebiet der semantischen Datenqualität:

(A) Ebene des Datenmodells

- Die Modellierung von Domänen der Realwelt in einer Ontologie und das Mapping dieser Ontologie auf das physische Modell eines Informationssystems tragen zu einer verbesserten Abstimmung zwischen Entitäten der Realwelt und Datenbankstrukturen bei.
- Ontologiekonzepte können gleichzeitig auf verschiedene Informationssysteme (mit heterogenen Datenbankschemata) abgebildet werden. Dies trägt zu einer besseren Abstimmung von Datenbankstrukturen verteilter betrieblicher Informationssysteme bei.

(B) Ebene des Datenbestands

- Technisch weniger versierte Benutzer können semantische Abfragen, die von Implementierungsdetails abstrahieren, erstellen (Abfragegenerierung).
- Der Nutzen von Ontologien zur Erfassung der Semantik der vorliegenden Datenbank erlaubt die Analyse der Datenqualität ohne sich mit technischen Fragen zu beschäftigen (Analyseunterstützung).

Eine konkrete weiterführende Anwendung unsers Beitrags liegt beispielsweise im Auditing von Informationssystemen bezüglich der Eignung zur Unterstützung neuer Geschäftsmodelle oder der Erfüllung der Anforderungen neuer gesetzlicher Regelungen. Durch die Modellierung von Konzepten, die heute vielleicht noch nicht benötigt werden, aber in Zukunft von Interesse sein werden, können eventuell in der Zukunft auftretende Datendefizite frühzeitig erkannt werden.

Beschränkungen

Die Ergebnisse in dieser Arbeit basieren auf einer begrenzten Anzahl von Fallstudien und Experteninterviews. Zudem erlaubt der interpretivistische Forschungsansatz keine allgemeingültigen Aussagen. Wir verstehen unseren Ansatz deshalb nur als Startpunkt auf dem Weg zu einem semantischen Datenqualitätsmanagement.

Zukünftige Forschung

Wir schlagen vor, das in diesem Artikel eingeführte Vorgehensmodell und das Werkzeug bei der Durchführung zukünftiger Forschung auf dem Gebiet der Datenqualität zu

verwenden. Außerdem sollte das Konzept zum Zwecke der Evaluation und Weiterentwicklung auf weitere Domänen ausgeweitet werden. Wir glauben, dass semantischen Aspekten des Datenqualitätsmanagement in Zukunft sowohl in Theorie als auch in der Praxis eine zunehmend wichtige Rolle zukommen wird.

5 Danksagung

Dieser Beitrag wurde durch die Förderung des BMBF-Projektes „ManKIP“ (Management kreativitätsintensiver Prozesse, Förderkennzeichen 01FM07061) im Rahmen des Förderprogramms „Hightech-Strategie für die moderne Arbeitswelt“ ermöglicht. Wir danken an dieser Stelle dem Projektträger Deutsches Zentrum Luft- und Raumfahrt (DLR) für die Unterstützung.

6 Literaturverzeichnis

- [Ag05] Agosta, L.: Trends in Data Quality. In: DM Review, 15 (2) 2005, S. 34-35.
- [BD07] Berson, A.; Dubov, L.: Master Data Management and Customer Data Integration for a Global Enterprise. McGraw-Hill, New York, 2007.
- [BGW08] Becker, J.; Glaser, J.; Winkelmann, A.: Auszug aus der aktuellen IT-Verbundgruppenstudie. In: Retail Technology Journal (03) 2008.
- [BJ07a] Becker, J.; Janiesch, C.; Pfeiffer, D.: Context-Based Modelling. In: Proceedings of the 11th Pacific Asia Conference on Information Systems (PACIS 2007), Auckland, New Zealand, 2007; S.143.
- [BJ07b] Becker, J.; Janiesch, C.; Pfeiffer, D.: Reuse Mechanisms in Situational Method Engineering. In: Proceedings of the IFIP WG 8.1 Working Conference on Situational Method Engineering, Geneva, Switzerland, 2007; S. 79-93.
- [BJ07c] Becker, J.; Janiesch, C.; Pfeiffer, D.: Towards More Reuse in Conceptual Modeling. In: Proceedings of the 19th International Conference on Advanced Information Systems Engineering (CaiSE 2007), Trondheim, Norway, 2007; S. 81-84.
- [BW08] Becker, J.; Winkelmann, A.: Handelscontrolling. Berlin, Heidelberg, New York, 2008.
- [BUV01] Becker, J.; Uhr, W.; Vering, O.: Retail Information Systems Based on SAP Products. Springer, Berlin u.a., 2001.
- [Ca04] Capgemini: Internal Data Alignment: Learning from Best Practices. Results of an Internal Data Alignment Survey, 2004.
- [CFP03] Cappiello, C.; Francalanci, C.; Pernici, B.: Time-Related Factors of Data Quality in Multichannel Information Systems. In: Journal of Management Information Systems, 20 (3) 2003; S. 71-91.
- [En99] English, L.P.: Improving Data Warehouse and Business Information Quality. Wiley, New York 1999.
- [GF95] Gruninger, M.; Fox, M.S.: Methodology for the Design and Evaluation of Ontologies, Montreal, 1995.

- [Gr93] Gruber, T.: A Translation Approach to Portable Ontology Specifications. In: Knowledge Management Acquisition, 5 (2) 1993; S. 199-220.
- [Gr05] Griffin, J.: The Master Data Challenge. In: DM Review, 15 (5) 2005; S. 85.
- [Gr08] Gruber, T.: Ontology. In (Liu, L.; Tamer Özsu, M. Hrsg.): Encyclopedia of Database Systems, Springer, 2008.
- [HKL95] Hirschheim, R.; Klein, H.K.; Lyytinen, K.: Information Systems Development and Data Modeling: Conceptual and Philosophical Foundations. Cambridge University Press, Cambridge, MA, 1995.
- [HLW99] Huang, K.-T.; Lee, Y.W.; Wang, R.Y.: Quality information and knowledge, Upper Saddle River, NJ, 1999.
- [HMP04] Hevner, A.R.; March, S.T.; Park, J.; Ram, S.: Design Science in Information Systems Research. In: MIS Quarterly, 28 (1) 2004, S. 75-105.
- [Ja91] Jackson, M.C.: Systems Methodology for the Management Sciences. Plenum, New York, 1991.
- [KPM04] Kahn, B.; Pierce, E.; Melkas, H.: IQ Research Directions. In (Chengalur-Smith, I.; Raschid, L.; Long, J.; Seko, C. Hrsg.): Proceedings of the 9th International Conference on Information Quality, Cambridge, MA, 2004; S. 326-332.
- [Ku04] Kuipers, P.: Data in Dire Need of a Spring Clear. In: Elsevier Food International, 3, S. 74-79.
- [Lo05] Loshin, D.: Master Data Standards and Data Exchange. In: DM Review, 15 (8) 2005; S. 72-77.
- [LSK02] Lee, Y.W.; Strong, D.M.; Kahn, B.K.; Wang, R.Y.: AIMQ: A Methodology for Information Quality Assessment. In: Information & Management, 40 (2) 2002; S. 133-146.
- [Mi97] Midgley, G.: Mixing methods: Developing Systemic Intervention. In (Mingers, J.; Gill, A., Hrsg.): Multimethodology: The Theory and Practice of Combining Management Science Methodologies, Wiley, Chichester, 1997.
- [NF05] Necib, C.B.; Freytag, J.-C.: Query Processing Using Ontologies. In: Proceedings of the Proceedings of the 17th Conference on Advanced Information Systems Engineering (CAISE'05), Porto, Portugal, 2005.
- [NM08] Noy, N.F.; McGuinness, D.L.: Ontology Development 101: A Guide to Creating Your First Ontology. In Stanford, CA, 2008.
- [PTR07] Peffers, K.; Tuunanen, T.; Rothenberger, M.A.; Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. In: MIS Quarterly, 24 (3) 2007, S. 45-77.
- [Pi04] Pierce, E.M.: Assessing Data Quality with Control Matrices. In: Communications of the ACM, 47 (2) 2004; S. 82-86.
- [PLW02] Pipino, L.L.; Lee, Y.W.; Wang, R.Y.: Data Quality Assessment. In: Communications of the ACM, 45 (4) 2002; S. 211-218.
- [PS05] Price, R.; Shanks, G.: A Semiotic Information Quality Framework: Development and Comparative Analysis. In: Journal of Information Technology, 20 (2) 2005; S. 88-102.
- [Pr 06] Prud'hommeux, Eric: SPASQL: SPARQL Support in MySQL. In: Proceedings of the XTech 2006, Amsterdam, The Netherlands, 2006.

- [Re96] Redman, T.C.: Data Quality in the Information Age, Artech House, Boston, MA.
- [Sc91] Schechter, D.: Critical Systems Thinking in the 1980s: A Connective Summary. In (Flood, R.L.; Jackson, M.C., Hrsg.): Critical Systems Thinking: Directed Readings, Wiley, Chichester, 1991.
- [Sh99] Shanks, G.: Semiotic Approach to Understanding Representation in Information Systems. In: Proceedings of the Proceedings of the Information Systems Foundations Workshop – Ontology, Semiotics and Practice, 1999.
- [So00] Sowa, J.F.: Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks Cole, Pacific Grove, CA, 2000.
- [UG96] Ushold, M.; Gruninger, M.: Ontologies: Principles, methods and applications. In: Knowledge Engineering Review, 11 (2) 1996.
- [Us96] Ushold, M.: Building Ontologies: Towards a Unified Methodology. In: Proceedings of the Proceedings of Expert Systems 1996, the 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems, Cambridge, 1996.
- [Wa98] Wang, R.Y.: A Product Perspective on Total Data Quality Management. In: Communications of the ACM, 41 (2) 1998; S. 58-65.
- [We97] Weber, R.: Ontological Foundations of Information Systems. Coopers and Lybrand, Melbourne, 1997.
- [Wi06] Winkelmann, A.: Integrated Couponing. A Process-Based Framework for In-Store Coupon Promotion Handling in Retail. Logos, Berlin, 2006.
- [WW96] Wand, Y.; Wang, R.Y.: Anchoring Data Quality Dimensions in Ontological Foundation. In: Communications of the ACM, 39 (11) 1996; S. 86-95.
- [WYP98] Wang, R.Y.; Yang, W.L.; Pipino, L.L.; Strong, D.M.: Manage Your Information as a Product. In: Sloan Management Review, 39 (4) 1998; S. 95-105.

Effektive Architekturgestaltung auf Basis einer Typologie für Datenintegrationsarchitekturen

Reinhard Jung

Lehrstuhl für Wirtschaftsinformatik und Betriebliche Kommunikationssysteme
Universität Duisburg-Essen, Institut für Informatik und Wirtschaftsinformatik
Universitätsstraße 9, DE-45141 Essen
reinhard.jung@icb.uni-due.de

Abstract: Die Datenintegration ist eine der zentralen Voraussetzungen für die Durchführung von Integrationsvorhaben, sei es unternehmensintern oder -übergreifend. Der Entwurf einer Datenintegrationsarchitektur ist deshalb von großer Bedeutung. Der Aufsatz identifiziert fünf Merkmale, mit deren Hilfe sich Datenintegrationsarchitekturen beschreiben und unterscheiden lassen. Durch Kombination aller miteinander verträglicher Merkmalsausprägungen entstehen 22 grundsätzlich mögliche Datenintegrationsarchitekturtypen. Eine Gegenüberstellung von fachlichen Anforderungen an Daten, die beispielsweise aus Geschäftsprozessen ableitbar sind, mit den Eigenschaften der Datenintegrationsarchitekturtypen ermöglicht eine Vorauswahl geeigneter (effektiver) Typen.

1 Einleitung

Die häufige Änderung von Geschäftsmodellen und Leistungsportfolios sowohl von Unternehmen als auch von Behörden zwingt zu hoher Anpassungsfähigkeit (Flexibilität) im Bereich der Anwendungssysteme und der IT-Architekturen. In realen Szenarien kann zu diesem Zweck nicht das gesamte Anwendungsportfolio ersetzt werden. Stattdessen sind neue Anwendungssysteme mit den bestehenden zu integrieren, um die Effektivität des betrieblichen Informationssystems aufrecht zu erhalten oder herzustellen. Ein zentraler Schritt eines solchen Integrationsvorhabens ist die Integration der Daten aller beteiligten Systeme (Datenintegration).

Prinzipiell steht eine ganze Reihe von Architekturkonzepten zur Verfügung, die für eine Integration grundsätzlich geeignet erscheinen. Beispiele sind so unterschiedliche Architekturkonzepte wie Data-Warehouse-Systeme, Operational Data Stores und serviceorientierte Architekturen. Darüber hinaus ist auch im Bereich der Technologien inzwischen eine große Menge an Optionen verfügbar, beispielsweise Middleware-Produkte und -Standards, Integrations- oder Message-Broker und Web Services. Viele Architekturkonzepte sind allerdings nicht universell einsetzbar. Ein Data-Warehouse-System beispielsweise ist – zumindest in der „klassischen“ Sichtweise – aufgrund der Charakteristika von Datenextraktions- und Konsolidierungsprozessen nicht in der Lage, Echtzeit-Daten bereitzustellen oder etwa die Modifikation von Data-Warehouse-Daten in die

operativen Anwendungssysteme zu propagieren (eine grundsätzlich andere Sichtweise wird vertreten bspw. in [Ze03]).

Bevor eine konkrete Architektur gestaltet und eine Entscheidung für bestimmte Technologien und Software-Lösungen gefällt werden kann, muss eine genaue Spezifikation der Anforderungen erfolgen. Der vorliegende Beitrag widmet sich der Frage, wie anhand dieser Anforderungen effektive Architekturen für die Datenintegration gestaltet werden können. Zu diesem Zweck wird, nach einer Festlegung der begrifflichen Grundlagen (Kapitel 2), eine Menge von Merkmalen vorgestellt, mit deren Hilfe sich Datenintegrationsarchitekturen beschreiben und unterscheiden lassen (Kapitel 3). In Kapitel 4 wird die Merkmalsmenge genutzt, um eine Typologie für Datenintegrationsarchitekturen aufzubauen; die Überprüfung der Typologie erfolgt durch Anwendung der Merkmale auf bekannte Datenintegrationsarchitekturen. Ein abstraktes Anwendungsbeispiel für die Klassifikation findet sich in Kapitel 5. Die Typologie wird dabei genutzt, um auf Basis von Anforderungen effektive, also mit Blick auf die Anforderungen geeignete Architekturen auswählen zu können. Wirtschaftlichkeitsüberlegungen (Effizienz), die u.a. die erforderliche Anpassung von Altsystemen erfassen, werden in diesem Aufsatz ausgelammert. Der Aufsatz schließt mit einer Zusammenfassung und einem Ausblick (Kapitel 6).

2 Begriffliche Grundlagen

Da Integration und insbesondere Datenintegration zu den traditionellen Analyse- und Gestaltungsgebieten der Informatik- und Wirtschaftsinformatik-Forschung gehören [Ro99], existiert eine Fülle von teilweise stark divergierenden Definitionen. Eine ausführliche Begriffsanalyse findet sich in [Ju06, 34 ff.]. Um zu tragfähigen Definitionen für diesen Beitrag zu gelangen, bietet es sich an, zunächst die in der aktuellen Literatur genannten Integrationsarten zu untersuchen:

- [RMB01, 17 ff.] unterscheiden Präsentations-, Daten- und Funktionsintegration.
- [FS01, 215 ff.] unterscheiden Daten-, Funktions- und Objektintegration.
- [Li00] unterscheidet neben Präsentations-, Daten- und Objektintegration auch die Integration von Applikationen durch Kopplung mit Hilfe von Schnittstellen.
- [BU02] unterscheiden Daten-, Applikations- und Prozessintegration, wobei Applikationsintegration als Kombination von Daten-, Funktions- und Präsentationsintegration aufgefasst wird.

Bereits die Unterscheidung von Integrationsarten zeigt ein heterogenes Bild. Als weitgehender Konsens in der Literatur lassen sich allerdings die Abhängigkeiten zwischen den verschiedenen Arten festhalten: Datenintegration ist eine Voraussetzung für die Funktionsintegration, und die Objektintegration ist eine alternative Sichtweise auf die Daten- und Funktionsintegration. Schließlich wird die Daten- und Funktionsintegration bzw. die Objektintegration als Voraussetzung für die Prozessintegration gesehen. Da der Auslöser von Datenintegrationsmaßnahmen primär in fachlichen Anforderungen zu sehen ist, liegt diesem Aufsatz entsprechend eine eher betriebswirtschaftliche Definition zugrunde:

Datenintegration ist der „Zustand, bei dem Aufgabenträger innerhalb eines Untersuchungsbereichs Zugriff auf die Informationsobjekte haben, die für die Aufgabenerfüllung erforderlich sind. Die Informationsobjekte müssen dabei den aufgaben- und aufgabenträgerspezifischen Qualitätsanforderungen genügen.“ [Ju06, 45]. Dabei ist ein Informationsobjekt eine zweckbezogene Zusammenfassung von miteinander in Beziehung stehenden Daten. Informationsobjekte können eine einfache Struktur aufweisen (z.B. Informationsobjekttyp „Kunde“) oder eine komplexe Struktur (z.B. Informationsobjekttyp „Umsatzvolumen nach Kunde, Region und Zeitpunkt“). Unsere Definition von Datenintegration schließt implizit alle denkbaren Verwendungen von Daten ein, also sowohl eine transaktionsorientierte und typischerweise modifizierende Nutzung als auch eine analytische, in der Regel nur lesende Nutzung. Ferner wird wie in [BU02, 416] davon ausgegangen, dass die Konsistenz des Gesamtdatenbestands zu erhalten ist.

Die bereits existierenden IKS werden zusammen als Applikationsarchitektur bezeichnet. Im Rahmen eines Integrationsvorhabens werden dieser Architektur zielgemäß (Ziel: Datenintegration) Integrationskomponenten hinzugefügt. Die Applikationsarchitektur bildet zusammen mit den Integrationskomponenten die Datenintegrationsarchitektur (DIA). Durch Abstraktion von konkreten Datenintegrationsarchitekturen erhält man über die Zwischenstufe von DIA-Modellen (vgl. dazu auch [Si04, 315]) schließlich DIA-Typen (vgl. dazu Abb. 1). Dabei richtet sich die Zusammenfassung zu einem DIA-Typ danach, inwieweit die betreffenden Datenintegrationsarchitekturen (oder auch DIA-Modelle) strukturelle Ähnlichkeit aufweisen.

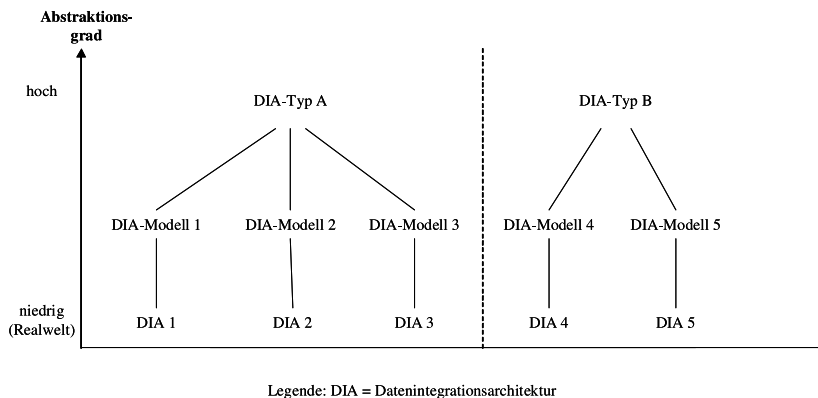


Abbildung 1: Abgrenzung der Begriffe „DIA“, „DIA-Modell“ und „DIA-Typ“ anhand eines abstrakten Beispiels

Zum Zweck einer begrifflichen Differenzierung werden Daten, die für neue Aufgaben benötigt werden und durch Integration erzeugt werden, als globale Daten bezeichnet. Die Daten der bereits vorhandenen Applikationen werden hingegen als lokale Daten bezeichnet. Die Begriffe entstammen dem Forschungsgebiet „Schemaintegration“ (vgl. z.B. [LNE89]).

Als Ergänzung zu den begrifflichen Grundlagen sei noch der Begriff Enterprise Application Integration (EAI) erwähnt, der hier thematische Relevanz besitzt, inzwischen aller-

dings zu Recht als Schlagwort bezeichnet wird [Ho03, 45]. Die Begriffsvielfalt sei anhand der folgenden Definitionen bekannter Vertreter der EAI-Szene aufgezeigt, die deutlich voneinander abweichen:

- Ruh et al. bezeichnen die Erstellung neuer Geschäftslösungen durch Kombination von (bestehenden) Applikationen unter Verwendung von Middleware als EAI [RMB01, 2].
- Linthicum definiert EAI als das uneingeschränkte Teilen von Daten und Geschäftsprozessen zwischen beteiligten Applikationen und Datenbanken innerhalb eines Unternehmens [Li00, 3].
- Stonebraker bezeichnet den Mechanismus zur Integration voneinander isolierter Informationen als EAI [St99, 2].

Der von uns vorgeschlagene Ansatz für den Entwurf der Datenintegration ist mit den genannten Interpretationen von EAI insoweit vereinbar, als er (a) bestehende Applikationen berücksichtigt und (b) auf die Integration („sharing“) von Daten zwischen Applikationen und Datenbanken gerichtet ist. Unser Ansatz entspricht damit weitgehend der EAI-Definition von Stonebraker. Wir sehen in der von uns angestrebten Datenintegration eine zentrale Voraussetzung für umfassendere Ansätze, die in den beiden anderen Definitionen implizit oder explizit zum Ausdruck kommen, insbesondere für die Prozessintegration.

3 Merkmale von Datenintegrationsarchitekturen

In diesem Abschnitt wird untersucht, mit Hilfe welcher Merkmale sich Datenintegrationsarchitekturen beschreiben und damit unterscheiden lassen; eine ausführliche Herleitung findet sich in [Ju06, 192 ff.].

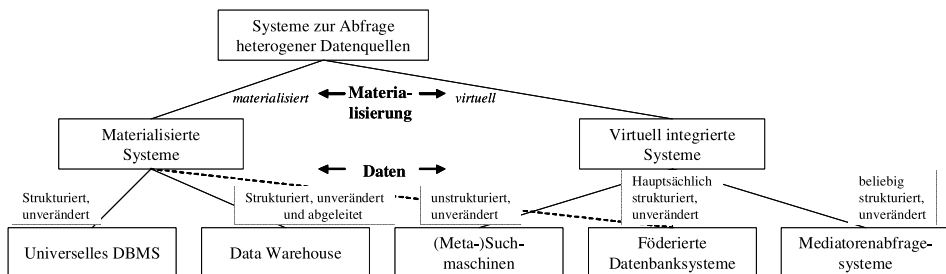


Abbildung 2: Klassifikation von Systemen zur Abfrage heterogener Datenquellen nach [DD99, 65]

Als Ausgangspunkt zur Identifikation von Merkmalen lässt sich die von [DD99] vorgeschlagene Klassifikation von Systemen zur Abfrage heterogener Datenquellen (vgl. Abb. 2) nutzen. Ein erstes Gliederungskriterium ist dort die Materialisierung der integrierten Daten mit den Ausprägungen „materialisiert“ und „virtuell“. Auf einer zweiten Stufe wird dann nach der Beschaffenheit der integrierten (globalen) Daten hinsichtlich Strukturierung und Modifikation gegenüber den lokalen Daten unterschieden. Obwohl dieses

zweite Merkmal für die Beschreibung einer Datenintegrationsarchitektur nicht wesentlich ist, führen die resultierenden Ausprägungen zur Identifikation eines weiteren wichtigen Merkmals, nämlich der Architekturtopologie, denn die auf der untersten Ebene in der Abbildung unterschiedenen Architekturen weichen in ihrer Topologie stark voneinander ab. In den beiden folgenden Abschnitten werden die Merkmale „Architekturtopologie“ und „Materialisierung“ dargestellt.

3.1 Architekturtopologie

Bezogen auf das Merkmal „Architekturtopologie“ lassen sich drei Formen unterscheiden: multilaterale Kopplung, Föderation und Fusion. Die Unterscheidung bezieht sich insbesondere darauf, wie zentral oder dezentral bestimmte Aufgaben innerhalb der Architektur organisiert sind. Zur Veranschaulichung dieser drei Formen wird im Folgenden von drei vorhandenen Applikationen (vA) und einer neuen Applikation (nA) ausgegangen, wobei die nA der Unterstützung neuer Aufgaben dient und damit der „Nachfrager“ von integrierten Daten ist.

Multilaterale Kopplung (Peer-to-peer)

Das entscheidende Merkmal der multilateralen Kopplung liegt darin, dass Komponenten eines betrachteten Systems mit allen übrigen Komponenten des Systems verbunden werden (können). Eng verbunden mit der Idee der Kopplung ist der Begriff „Peer-to-peer“, der im Sinne eines Netzwerks aus Komponenten (Peers) wie folgt charakterisiert wird [SF02, 587]:

- Jede Komponente kann Server- und Client-Funktionalität aufweisen, d.h. Anfragen sowohl stellen als auch beantworten.
- Die Kommunikation zwischen den Komponenten erfolgt direkt, d.h. ohne Zwischenschaltung einer zentralen Instanz.
- Die Komponenten behalten vollständig ihre Autonomie.

Im vorliegenden Kontext bringt es die Autonomie der Applikationen und das Fehlen einer zusätzlichen Komponente mit sich, dass Erweiterungen oder Modifikationen der zu koppelnden vA erforderlich werden. So ist insbesondere die Transaktionsverwaltung eine zusätzlich durchzuführende Aufgabe. Dabei ist festzulegen, ob die vA die Transaktionen auf dem eigenen Datenbestand überwachen, oder ob diejenige Applikation die Überwachungsfunktion übernimmt, welche die Transaktion ausgelöst hat. Eine weitere Aufgabe bei der Integration der zu koppelnden Applikationen ist die Datenformatkonvertierung.

Föderation

Der Begriff „Föderation“ entstammt dem Bereich der Datenbanktechnik. Dort wird ein Multidatenbanksystem, dessen Komponentensysteme (Datenbanken) eine gewisse Selbständigkeit (Autonomie) behalten, als föderiertes Datenbanksystem bezeichnet [Co97, 41]. Entsprechend lässt sich im Kontext dieses Beitrags der Begriff „Föderation“

verwenden. Unter einer Föderation wird eine Architekturtopologie verstanden, in der bestimmte Aufgaben im Zusammenhang mit der Datenintegration (z.B. Transaktionsverwaltung, Datenformatkonvertierung) von einer zentralen, koordinierenden Komponente ausgeführt werden. Gleichzeitig behalten die an der Föderation beteiligten Komponenten ihre Autonomie.

Fusion

Bei der Fusion existiert eine zentrale Komponente, welche die nA bedient und dabei eine vollständige Kontrolle über die involvierten Datenbestände ausübt; insbesondere die Transaktionskontrolle wird dabei zentralisiert. Dies kann durch zwei Varianten erreicht werden:

- Faktische Fusion: Die lokalen Daten werden in eine zentrale Datenbasis migriert, die der koordinierenden Komponente zugeordnet ist. Nachfolgend werden die vA an die veränderte Datenhaltung angepasst oder – falls ihre Funktionalität anderweitig abgedeckt wird – ausser Betrieb gesetzt.
- Virtuelle Fusion: Die vA werden in einen Verbund unter zentraler Kontrolle durch die koordinierende Komponente eingebunden, wobei die vA ihre Autonomie vollständig verlieren. Auch hier ist ein globales Schema zu erstellen.

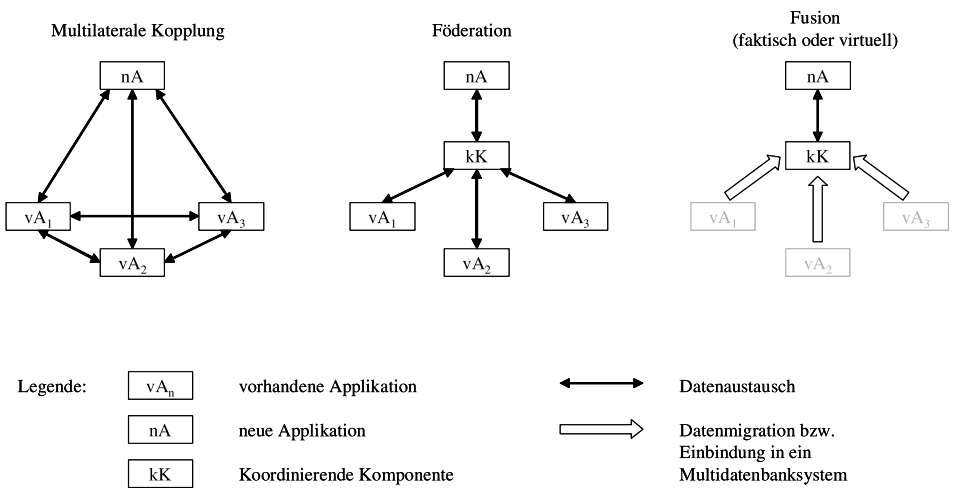


Abbildung 3: Architekturtopologien

In Abb. 3 sind die drei möglichen Architekturtopologien veranschaulicht. Bei der dritten Variante sind die vA schattiert dargestellt, weil sie bei der faktischen Fusion ggfs. entfallen.

3.2 Materialisierung

Das Merkmal „Materialisierung“ bezieht sich auf die Frage, ob die zur Deckung des Informationsbedarfs erforderlichen Daten in einer integrierten Form vorgehalten werden (materialisiert, „fetch in advance“) oder aber erst zum Zeitpunkt der Anfrage (virtuell, „fetch on demand“) konsolidiert und weitergegeben werden [LB02, 76; SH01, 556].

Materialisierung ist bei den Architekturtopologien „Föderation“ und „multilaterale Kopplung“ möglich. Im ersten Fall beinhaltet die koordinierende Komponente eine zentrale Datenbasis, welche die replizierten Daten aufnimmt. Im zweiten Fall werden bei jeder Applikation zusätzlich zu den „eigenen“ Daten Replikate von Daten anderer Applikationen abgelegt.

Eine Materialisierung bringt Vorteile mit sich, da ein Datenzugriff dann unabhängig von einer eingeschränkten Verfügbarkeit der Kommunikationsverbindungen und der Applikation erfolgen kann, welche die lokalen Daten führt. Gleichzeitig entstehen aber auch Nachteile [NI02, 431], beispielsweise bei der Konsistenzsicherung zwischen Replikaten bei Nicht-Verfügbarkeit der beteiligten Komponenten.

Ein Spezialfall ergibt sich, wenn bestimmte Daten in eine separate Datenbank migriert werden. In [Br01, 220 f.] wird dieser Fall als „Shared data solution“ bezeichnet: Daten, die von mehreren Applikationen benötigt werden, werden in eine zentrale Datenbank migriert; die übrigen (lokalen) Daten verbleiben bei den jeweiligen Applikationen. Da mit der zentralen Datenbank auch eine koordinierende Komponente in Form eines zugehörigen Datenbankmanagementsystems erforderlich wird, ist diese Variante der Architekturtopologie „Föderation“ zuzurechnen.

Es ergeben sich die folgenden Merkmalsausprägungen für die Materialisierung: „materialisiert“, „virtuell“ und „teilweise migriert“. Die Ausprägung „teilweise migriert“ repräsentiert den von Britton beschriebenen Ansatz.

3.3 Transaktionstyp

Einen Hinweis auf ein weiteres Merkmal, das zur Unterscheidung von Architekturen für die Datenintegration geeignet ist, liefert [St99] im Zusammenhang mit der Klassifikation von EAI-Technologien. Stonebraker unterscheidet u.a. danach, welcher Zweck mit dem Technologieeinsatz verfolgt wird. Zum einen wird die Aktualisierung von Daten und zum anderen das Lesen genannt.

Als Merkmal einer Datenintegrationsarchitektur wird in diesem Zusammenhang im Folgenden von dem Transaktionstyp gesprochen, der innerhalb der Architektur auf globalen Daten zulässig ist. Unterschieden werden typischerweise – wie auch bei verteilten Datenbanksystemen [CA98, 280, Hu97, 52 ff.] – Architekturen, die ausschließlich lesende Zugriffe auf globale Daten zulassen, und solche, die zusätzlich auch schreibende Zugriffe auf diese Daten ermöglichen. Ein schreibender Zugriff ist dabei ein Zugriff, der Daten erzeugt, verändert oder löscht. Sofern schreibende Zugriffe auf globale Daten

zugelassen werden, ergibt sich implizit die Anforderungen, die zugehörigen lokalen Daten zu aktualisieren (vgl. dazu den folgenden Abschnitt).

3.4 Aktualisierungskontrolle

Zwischen lokalen Daten und dem Ort des Datenzugriffs (globale Daten), bei dem es sich im Fall der Materialisierung um eine dedizierte Datenbank handeln kann, ergeben sich zwei Aktualisierungsperspektiven:

- *Aktualisierungskontrolle lokal-zu-global:* Nach der initialen Materialisierung von globalen Daten kommt es häufig vor, dass die zugrunde liegenden lokalen Daten modifiziert werden. Es stellt sich daher die Frage, nach welchen Regeln und in welchem Zeitrahmen die Aktualisierung zwischen lokalen und globalen Daten erfolgen soll. Mit Blick auf den Modifikationszeitpunkt der lokalen Daten werden die beiden folgenden Varianten unterschieden: sofortige (synchrone) und verzögerte/spätere (asynchrone) Aktualisierung der globalen Daten.
- *Aktualisierungskontrolle global-zu-lokal:* Bei dieser Aktualisierungsperspektive wird davon ausgegangen, dass globale Daten (virtuelle oder materialisierte) verändert wurden. Auch hier stellt sich die Frage, wie die Aktualisierung der lokalen Daten entsprechend der Modifikation globaler Daten erfolgen soll; es ergeben sich wiederum die oben beschriebenen Möglichkeiten einer synchronen oder asynchronen Aktualisierung.

4 Datenintegrationsarchitekturtypen

Im vorhergehenden Abschnitt wurden Architekturtopologie, Materialisierung, Transaktionstyp und Aktualisierungskontrolle (lokal-zu-global, global-zu-lokal) als Merkmale vorgestellt, mit denen sich Datenintegrationsarchitekturen beschreiben lassen.

4.1 Typisierung

Um DIA-Typen identifizieren zu können, sind die Konstruktionsmerkmale zunächst in eine Reihenfolge zu bringen, welche die wechselseitigen Abhängigkeiten berücksichtigt:

- Die Global-zu-lokal-Aktualisierung ist nur relevant, wenn auf die (materialisierten oder virtuellen) globalen Daten auch schreibend zugegriffen wird.
- Die Lokal-zu-global-Aktualisierung ist nur dann erforderlich, wenn die globalen Daten materialisiert vorliegen.
- Wird als Architekturtopologie die Fusion gewählt, dann sind das Merkmal „Materialisierung“ sowie die davon abhängigen Merkmale irrelevant.

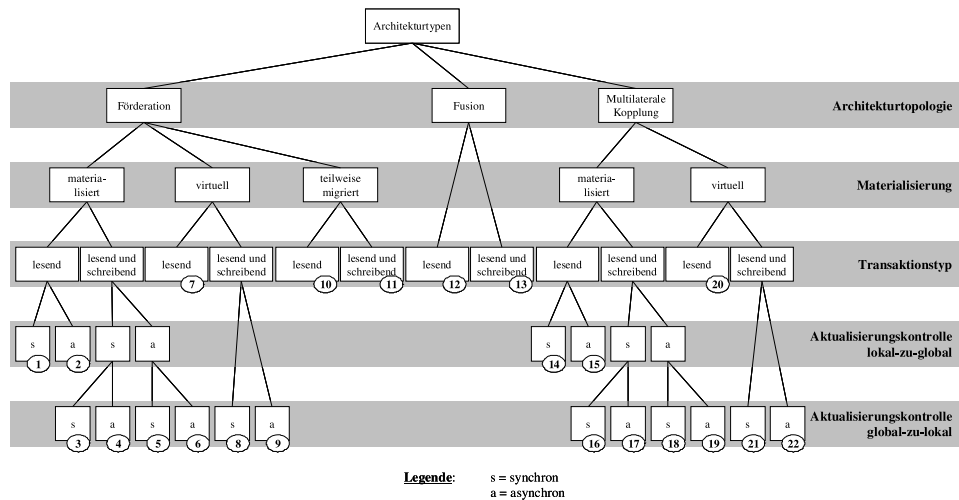


Abbildung 4: Merkmalsbaum für Datenintegrationsarchitekturen [Ju06, 222]

Aus den Abhängigkeiten lässt sich eine baumförmige Strukturierung der Merkmalsausprägungen ableiten. Innerhalb dieser Baumstruktur ist jede Merkmalsausprägung auf einer Ebene mit den möglichen Merkmalsausprägungen der darunter liegenden Ebene verbunden (vgl. Abb. 4).

Durch diese Anordnung entstehen Pfade von der Wurzel des Baums bis zu den verschiedenen Blattelementen, so dass jeder Pfad einen von insgesamt 22 möglichen DIA-Typen repräsentiert.

4.2 Überprüfung der Datenintegrationsarchitekturtypen

Die Überprüfung der Merkmalsmenge und damit der hergeleiteten DIA-Typen umfasst zwei Aspekte. Es ist zu untersuchen, ob sich die Merkmalsmenge einerseits vollständig und andererseits disjunkt verhält. Da sich Vollständigkeit in diesem Kontext nicht beweisen lässt, unternehmen wir den Versuch, die Merkmalsmenge zu falsifizieren. Zu diesem Zweck erfolgt eine Beschreibung der aus Literatur und Praxis bekannten Datenintegrationsarchitekturen mit Hilfe der Merkmalsmenge. Auf diese Weise kann auch die Disjunktheit überprüft werden: Unterschiedliche Datenintegrationsarchitekturen sollten unterschiedliche Kombinationen von Merkmalsausprägungen aufweisen. Einen weiteren Überprüfungsschritt stellt die Einordnung anderer Klassifikationsansätze in unseren Ansatz dar (vgl. Abschnitt 6).

Data-Warehouse-Systeme

Die Konsolidierung von Daten aus unterschiedlichen betrieblichen Funktionsbereichen und ihre Weitergabe in – bezogen auf die Hierarchie – vertikaler Richtung und damit für dispositive Zwecke stellt eine wesentliche Facette der Datenintegration dar [PRW01, 181]. Aus verschiedenen Gründen ist es allerdings nicht möglich bzw. nicht praktikabel,

„auf“ den operativen Applikationen eine Datenintegration zu realisieren, die lediglich „virtuell“ ist und bei Bedarf ad hoc erzeugt wird. Beispiele sind eine für dispositive Zwecke zu „schmale“ Datenhistorie in den operativen Applikationen und mögliche Laufzeitengpässe, die durch umfangreiche Abfragen auf den operativen Daten entstehen könnten.

Der typische Lösungsansatz in diesem Kontext ist ein Data-Warehouse-System (DW), also eine themenorientierte, integrierte, zeitorientierte und nicht-volatile Datenbasis, die zur Unterstützung von Managemententscheidungen eingesetzt werden kann [In92]. Die in einem DW verfügbaren Daten weisen einige besondere Eigenschaften auf:

- *Redundanz*: Die Daten sind Kopien von lokalen Daten oder aus diesen abgeleitete Daten.
- *Stabilität des Datenbestands in definierten Zeitintervallen*: Da die Daten in regelmäßigen Abständen aus den operativen Datenbanken in das DW kopiert (extrahiert) werden, ist der Datenbestand im DW in den Zeitintervallen zwischen zwei Extraktionszeitpunkten stabil.
- *Eingeschränkte Datenaktualität*: In den vorgenannten Zeitintervallen nimmt die Aktualität der Daten – verglichen mit den lokalen Daten – ab.
- *Informationscharakter der Daten*: Der Datenfluss innerhalb eines DW ist unidirektional, d.h., die Daten besitzen lediglich Informationscharakter. Ihre Modifikation mit dem Ziel, damit durch Propagierung auch die lokalen Daten zu verändern, ist grundsätzlich nicht vorgesehen.

In Abb. 5 ist die Klassifikation eines DW mit Hilfe eines morphologischen Kastens dargestellt.

Architekturtopologie	Förderung		Fusion	Multilaterale Kopplung
Materialisierung	materialisiert	virtuell	teilweise migriert	nicht anwendbar
Transaktionstyp	lesend		lesend und schreibend	
Aktualisierungskontrolle lokal-zu-global	synchron	asynchron		nicht anwendbar
Aktualisierungskontrolle global-zu-lokal	synchron	asynchron		nicht anwendbar

Abbildung 5: Morphologischer Kasten für ein Data-Warehouse-System (DIA-Typ 2)

Operational Data Stores

Wie auch das DW weist ein Operational Data Store (ODS) eine dedizierte Datenbank auf, die lokale Daten in replizierter Form enthält. Das Ziel ist hier allerdings nicht die Unterstützung von Entscheidungsprozessen, sondern die Bereitstellung von integrierten Daten für operative Geschäftsprozesse, beispielsweise für ein Kundenselbstbedienungssystem (z.B. Online-Banking). Aus dieser Zielsetzung resultieren von einem DW abwei-

chende Anforderungen und spezifische technische Eigenschaften des ODS [Wi00, 135 f.]:

- Ein ODS beinhaltet in der Regel keine historisierten, sondern integrierte aktuelle Daten. Die Daten sind mit Blick auf operative Geschäftsprozesse zu Datenobjekten zusammengefasst, d.h. ihre Struktur ist auf die Verwendung in Transaktionen ausgerichtet.
- Neben dem lesenden Zugriff bedingen Transaktionen in Geschäftsprozessen in der Regel auch einen schreibenden Zugriff auf die integrierten (globalen) Daten, mit der Konsequenz, dass die lokalen Daten möglichst umgehend synchronisiert (aktualisiert) werden müssen.
- Da die Daten in einem ODS aufgrund des Verwendungszwecks wesentlich aktueller als die in einem DW sein müssen, ist eine Aktualisierung in längeren Zeitintervallen nicht sinnvoll. Stattdessen sind die globalen Daten möglichst umgehend zu aktualisieren, wenn sich die lokalen Daten ändern.

Abb. 6 zeigt die Klassifikation eines ODS.

Architekturtopologie	Föderation		Fusion		Multilaterale Kopplung
Materialisierung	materialisiert	virtuell	teilweise migriert	nicht anwendbar	
Transaktionstyp	lesend			lesend und schreibend	
Aktualisierungskontrolle lokal-zu-global	synchron		asynchron		nicht anwendbar
Aktualisierungskontrolle global-zu-lokal	synchron		asynchron		nicht anwendbar

Abbildung 6: Morphologischer Kasten für einen Operational Data Store (DIA-Typen 3 bis 6)

Enterprise-Resource-Planning-Systeme

Architekturtopologie	Föderation		Fusion		Multilaterale Kopplung
Materialisierung	materialisiert	virtuell	teilweise migriert	nicht anwendbar	
Transaktionstyp	lesend			lesend und schreibend	
Aktualisierungskontrolle lokal-zu-global	synchron		asynchron		nicht anwendbar
Aktualisierungskontrolle global-zu-lokal	synchron		asynchron		nicht anwendbar

Abbildung 7: Morphologischer Kasten für ein Enterprise-Resource-Planning-System (DIA-Typ 13)

Föderierte Datenbanksysteme

Föderierte Datenbanksysteme (FDBS) [HM79] gehören zu den so genannten Multidatenbanksystemen; ein FDBS ist also ein Zusammenschluss mehrerer Datenbanksysteme [SL90]. Ein charakteristisches Merkmal dabei ist, dass die Komponenten(datenbank)systeme eines FDBS weitgehend ihre Selbstständigkeit (Autonomie) behalten [Co97, 41]. In diesem Zusammenhang sind insbesondere die Ausführungs- und die Kommunikationsautonomie zu nennen [BBE99, 5]. Ausführungsautonomie besagt, dass die Komponentensysteme selbst „entscheiden“, wann sie Aufträge von anderen Komponenten ausführen. Kommunikationsautonomie besagt analog, dass sie selbst entscheiden, wann sie mit anderen Komponenten kommunizieren.

Die zentrale Komponente eines FDBS ist der so genannte *Föderierungsdienst*. Seine Aufgaben lassen sich grob wie folgt beschreiben:

- Auswertung und Weiterleitung von Anfragen der globalen Applikation(en) an die beteiligten Datenbankmanagementsysteme sowie Konsolidierung der Anfrageergebnisse und Rücksendung an die globalen Applikation(en).
- Auswertung von Schreibzugriffen der globalen Applikation(en) und Durchführung der entsprechenden Transaktionsverwaltung, was unter der Prämisse der Autonomieerhaltung bei den Komponenten und mit Blick auf verteilte Transaktionen durchaus zu Schwierigkeiten führen kann.

Ein FDBS ist somit ein datenintegrierendes System, denn der Föderierungsdienst (als einzige ergänzte Komponente) realisiert für globale Applikationen, also solche mit Datenintegrationsbedarf, „eine eigene Datenbankfunktionalität [...], so dass sich das Gesamtsystem dem globalen Nutzer, der über den Föderierungsdienst auf die Komponentensysteme zugreift, als ein Datenbanksystem darstellt“ [Co97, 50]. Ein FDBS weist bezüglich der integrierten Daten spezifische Eigenschaften auf:

- *Redundanzfreiheit*: Ein FDBS stellt lediglich einen Dienst „auf“ vorhandenen lokalen Daten zur Verfügung; die angeforderten Daten werden ad hoc integriert und weitergeleitet, sodass keine Redundanz durch Materialisierung entsteht.
- *Hohe Datenaktualität*: Sofern eine Anfrage sowie die spätere Auswertung und Weiterleitung des Ergebnisses durch den Föderierungsdienst in einem sehr kurzen Zeitraum ausgeführt werden kann, ist von einer – verglichen mit der Datenaktualität in den operativen Datenbanken – ähnlich hohen Datenaktualität auszugehen.
- *Eingeschränkter Zugriff*: Da die Komponenten-Datenbanksysteme eines FDBS ihre Autonomie in der Regel behalten, kann es dazu kommen, dass Datenzugriffe nicht immer (sofort) ausgeführt werden.

- *Schreibende Transaktionen:* Im Gegensatz zu einem DW sind bei einem FDBS auch schreibende Transaktionen auf globalen Daten mit Propagierung in die lokalen Datenbestände vorgesehen.

In Abbildung 8 ist die Klassifikation eines FDBS dargestellt.

Architekturtopologie	Föderation		Fusion	Multilaterale Kopplung
Materialisierung	materialisiert	virtuell	teilweise migriert	nicht anwendbar
Transaktionstyp	lesend		lesend und schreibend	
Aktualisierungskontrolle lokal-zu-global	synchron		asynchron	nicht anwendbar
Aktualisierungskontrolle global-zu-lokal	synchron		asynchron	nicht anwendbar

Abbildung 8: Morphologischer Kasten für ein föderiertes Datenbanksystem (DIA-Typen 7 bis 9)

5 Beispielhafte Anwendung

In diesem Abschnitt wird anhand eines abstrakten Beispiels aufgezeigt, wie die Klassifikation von Datenintegrationsarchitekturen verwendet werden kann, um auf Basis fachlicher Anforderungen zu einer Entscheidung hinsichtlich effektiver Datenintegrationsarchitekturen zu gelangen. In [Ju06, 212 ff.] wurden Merkmale des Informationsbedarfs aus Integrationssicht hergeleitet und daraufhin untersucht, inwiefern sie durch strukturelle Eigenschaften einer Datenintegrationsarchitektur beeinflusst werden (können). Zur Veranschaulichung wird hier auf eine Teilmenge der Merkmale zurückgegriffen (vgl. auch [Ju05]):

- *Antwortzeit:* Die Antwortzeit ist ein Maß für die maximale Zeit, die zwischen Anforderung und Bereitstellung von Daten vergehen darf.
- *Aktualität:* Die Aktualität drückt aus, inwieweit Daten den gegenwärtigen Zustand des zugrunde liegenden Realweltobjekts bzw. der zugrunde liegenden Realweltobjekte beschreiben sollen. In vielen Publikationen werden für besonders restriktive Anforderungen Ausprägungen wie „real-time“ und „near real-time“ verwendet.
- *Vollständigkeit:* Die Vollständigkeit drückt aus, inwieweit bei Datenbereitstellungen fehlende Werte akzeptabel sind. Die Forderung nach einer maximalen Vollständigkeit bei einer Verkaufsstatistik auf Basis von Monaten würde beispielsweise bedeuten, dass der Zugriff auf alle Monatsumsätze in einer bestimmten Periode erforderlich ist.
- *Verwendungsform:* Grundsätzlich sind zwei Verwendungsformen von Daten unterscheidbar. Zum einen können Daten für rein informative Zwecke eingesetzt werden, zum anderen ist es in manchen Szenarien erforderlich, Daten auch

verändern zu können. Im ersten Fall sprechen wir von der Verwendungsform „lesend“, im zweiten von „lesen und schreibend“.

Im Beispiel gehen wir davon aus, dass minimale Antwortzeiten sowie eine hohe Aktualität und Vollständigkeit der Daten gefordert werden sowie eine „lesende und schreibende“ Verwendung vorgesehen ist. Die folgenden Ausführungen beziehen sich auf Abb. Die Architekturtypen 10 und 12 sind nur von theoretischer Bedeutung und werden deshalb ausgeklammert. Bei beiden Typen ist eine Modifikation der integrierten Daten, die nur noch zentralisiert vorliegen, nicht möglich, sodass eine Einsetzbarkeit in der Praxis nicht gegeben ist.

Die vorgesehene Verwendungsform im Beispiel schränkt die Menge der DIA-Typen zunächst auf die Architekturtypen 3 bis 6, 8, 9, 11, 13, 16 bis 19, 21 und 22 ein, da nur diese für das Merkmal Transaktionstyp die Ausprägung „lesend und schreibend“ aufweisen. Die geforderte hohe Vollständigkeit schließt weiterhin Architekturtypen aus, die ausschließlich auf vorhandenen Datenressourcen aufsetzen und folglich Daten nur im Rahmen des ursprünglichen Verwendungszwecks vorhalten. Damit scheiden die Architekturtypen aus der Menge aus, die für das Merkmal „Materialisierung“ die Ausprägung „virtuell“ aufweisen, sodass nur noch die Architekturtypen 3 bis 6, 11, 13, 16 bis 19 in Frage kommen. In einem weiteren Selektionsschritt führt die geforderte hohe Aktualität bei den Architekturtypen mit replizierten Daten (Merkmal „Materialisierung“, Ausprägung „materialisiert“) zum Ausschluss aller Architekturtypen, die bei der „Aktualisierungskontrolle lokal-zu-global“ die Ausprägung „asynchron“ aufweisen. Das schränkt die Menge an noch zu überprüfenden Architekturtypen auf 3, 4, 11, 13, 16 und 17 ein. In einem letzten Schritt sind die Auswirkungen der geforderten kurzen Antwortzeit zu bedenken. Da die sechs in der Auswahl verbliebenen Architekturtypen über zusätzliche Datenbanken (mit globalen Daten) verfügen, ist die Antwortzeit bei lesenden Zugriffen selbst bei Nicht-Verfügbarkeit der lokalen Daten sicherlich als kurz einzustufen. Bei schreibenden Zugriffen auf die globalen Daten sind allerdings Aktualisierungen der lokalen Daten durchzuführen. Mit Blick auf die systemübergreifende Datenkonsistenz ist zu fordern, dass auch die hier durchzuführende Aktualisierung in Richtung der lokalen Datenbestände synchron erfolgt (Merkmal „Aktualisierungskontrolle global-zu-lokal“, Ausprägung „synchron“). Dadurch erfolgt ein letzter Eliminierungsschritt, als dessen Ergebnis lediglich die Architekturtypen 3, 11, 13 und 16 als effektiv (d.h. die Anforderungen erfüllend) einzustufen sind.

6 Zusammenfassung und Ausblick

Eine zentrale Voraussetzung für den zielgerichteten Entwurf von Architekturen ist die Möglichkeit, den Lösungsraum (hier: DIA-Typen) beschreiben zu können. Zu diesem Zweck wurde eine Menge von Merkmalen und deren Abhängigkeiten vorgestellt. Das Ergebnis sind 22 DIA-Typen, die jeweils spezifische Eigenschaften aufweisen. Eine Gegenüberstellung von fachlichen Anforderungen an Daten, die beispielsweise aus Geschäftsprozessen ableitbar sind, mit den Eigenschaften der DIA-Typen ermöglicht effektive (sachziel-konforme) Entwurfsentscheidungen (vgl. das abstrakte Beispiel in Kapitel 5). Die Validität der Merkmalsmenge sowie des Architekturselektionsprozesses aus dem

vorhergehenden Abschnitt konnte in drei Fallstudien aus dem Finanzdienstleistungsbereich bestätigt werden (vgl. [Ju06, 247 ff.]); in dieser Richtung erscheint eine weitere Überprüfung – auch in anderen Branchen – sinnvoll. Darüber hinaus ist zu überprüfen, ob die zusätzliche Berücksichtigung neuer Technologien und Konzepte, z.B. Data Streams, eine Modifikation der Merkmalsmenge und in der Folge der DIA-Typen bedingt. Einen weiteren Schwerpunkt zukünftiger Forschungsarbeit muss die Betrachtung der Effizienz bzw. der jeweils zu erwartenden Kosten (Formalziel) bilden, weil in einem Szenario mit mehreren als effektiv eingestuften DIA-Typen eine Entwurfsentscheidung auf Basis von Wirtschaftlichkeitsüberlegungen erfolgen muss.

7 Literaturverzeichnis

- [BBE99] Bouguettaya, A.; Benatallah, B.; Elmagarmid, A.: An Overview of Multidatabase Systems: Past and Present. In (Elmagarmid, A.; Rusinkiewicz, M.; Sheth, A. Hrsg.): *Heterogeneous Autonomous Database Systems*. Morgan Kaufman, San Francisco, 1999; S. 1-32.
- [Br01] Britton, C.: *IT Architectures and Middleware*. Addison-Wesley, Boston et al., 2001.
- [Bu02] Bunjes, B.; Friebe, J.; Götze, R.; Harren, A.: Integration von Daten, Anwendungen und Prozessen am Beispiel des Telekommunikationsunternehmens EWE TEL. In: *WIRTSCHAFTSINFORMATIK* 44 (2002) 5; S. 415-423.
- [CA98] Calvanese, D.; De Giacomo, G.; Lenzerini, M.; Nardi, D.; Rosati, R.: Information Integration: Conceptual Modeling and Reasoning Support. In (Halper, M. Hrsg.): *Proceedings of the 6th International Conference on Cooperative Information Systems (CoopIS'98)*, IEEE Computer Society, New York, 1998; S. 280-291.
- [Co97] Conrad, S.: *Föderierte Datenbanksysteme*. Springer, Berlin et al., 1997.
- [DD99] Domenig, R.; Dittrich, K.R.: An Overview and Classification of Mediated Query Systems. In: *ACM SIGMOD Record* 28 (1999) 3; S. 63-72.
- [FS01] Ferstl, O.K.; Sinz, E.J.: *Grundlagen der Wirtschaftsinformatik*, Band 1, 4. Aufl. München, Wien, Oldenbourg, 2001.
- [HM79] Hammer, M.; McLeod, D.: *On Database Management System Architecture*, Technical Report MIT/LCS/TM-141. Massachusetts Institute of Technology, Cambridge, 1979.
- [Ho03] Holten, R.: Integration von Informationssystemen. *WIRTSCHAFTSINFORMATIK* 45 (2003) 1; S. 41-52.
- [Hu97] Hull, R.: Managing Semantic Heterogeneity in Databases: A Theoretical Perspective. In (Mendelzon, A.; Özsoyoglu, Z.M. Hrsg.): *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*. ACM, New York, 1997; S. 51-61.
- [In92] Inmon, W.H.: *Building the Data Warehouse*. Wiley, New York et al., 1992.
- [Ju05] Jung, R.: Anforderungen an Informationsobjekttypen als Basis von Architekturentscheidungen bei der Datenintegration. In (Lenz, R.; Hasenkamp, U.; Hasselbring, W.; Reichert, M. Hrsg.): *EAI-Workshop 2005 – Enterprise Application Integration*. GITO-Verlag, Berlin, 2005; S. 82-89.

- [Ju06] Jung, R.: Architekturen zur Datenintegration – Gestaltungsempfehlungen auf der Basis fachkonzeptueller Anforderungen. DUV, Wiesbaden, 2006.
- [LB02] Lehner, W.; Bauer, A.: Data-Warehouse-Systeme – derzeitiger Stand und aktuelle Entwicklungen. In: Datenbank-Spektrum 2 (2002) 4; S. 76-78.
- [Li00] Linthicum, D.S.: Enterprise Application Integration. Addison-Wesley, Reading 2000.
- [LNE89] Larson, J.A.; Navathe, S.B.; Elmasri, R.: A Theory of Attribute Equivalence in Databases with Application to Schema Integration. In: IEEE Transactions on Software Engineering 15 (1989) 4; S. 449-463.
- [NI02] Niemann, H.; Hasselbring, W.; Wendt, T.; Winter, A.; Meierhofer, M.: Kopplungsstrategien für Anwendungssysteme im Krankenhaus. In: WIRTSCHAFTSINFORMATIK 44 (2002) 5; S. 425-434.
- [PRW01] Picot, A.; Reichwald, R.; Wigand, R.T.: Die grenzenlose Unternehmung, 4. Aufl. Gabler, Wiesbaden, 2001.
- [RMB01] Ruh, W.A.; Maginnis, F.X.; Brown, W.J.: Enterprise Application Integration – A Wiley Tech Brief. Wiley, New York et al., 2001.
- [Ro99] Rosemann, M.: Gegenstand und Aufgaben des Integrationsmanagements. In (Scheer, A.-W.; Rosemann, M.; Schütte, R. Hrsg.): Integrationsmanagement, Arbeitsbericht Nr. 65; Institut für Wirtschaftsinformatik der Westf. Wilhelms-Universität Münster, Münster, 1999, S. 5-18.
- [SF02] Schoder, D.; Fischbach, K.: Peer-to-Peer. In: WIRTSCHAFTSINFORMATIK 44 (2002) 6; S. 587-589.
- [SH01] Stonebraker, M.; Hellerstein, J.M.: Content Integration for E-Business. In (Sellis, T. Hrsg.): Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data. ACM Press, New York, 2001; S. 552-560.
- [Si04] Sinz, E.J.: Unternehmensarchitekturen in der Praxis – Architekturdesign am Reißbrett vs. situationsbedingte Realisierung von Informationssystemen. In: WIRTSCHAFTSINFORMATIK 46 (2004) 4; S. 315-316.
- [SL90] Sheth, A.; Larson, J.: Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases. In: ACM Computing Surveys 22 (1990) 3; S. 183-236.
- [St99] Stonebraker, M.: Integrating Islands of Information. In: EAI Journal (1999) September/October; S. 1-5.
- [Wi00] Winter, R.: Zur Positionierung und Weiterentwicklung des Data Warehousing in der betrieblichen Applikationsarchitektur. In (Jung, R.; Winter, R. Hrsg.): Data Warehousing Strategie. Springer, Berlin et al., 2000; S. 127-139.
- [Ze03] Zeh, T.: Data Warehousing als Organisationskonzept des Datenmanagements – Eine kritische Betrachtung der Data-Warehouse-Definition von Inmon. Informatik – Forschung und Entwicklung 18 (2003); S. 32-38.

Management von Business Intelligence Services

Thomas A. Horakh, Henning Baars, Hans-Georg Kemper

Lehrstuhl für ABWL und Wirtschaftsinformatik I
Universität Stuttgart
Breitscheidstr. 2c
70174 Stuttgart
{horakh | baars | kemper}@wi.uni-stuttgart.de

Abstract: Integrierte Ansätze zur Entscheidungsunterstützung – Business Intelligence (BI) – haben zwischenzeitlich eine große Verbreitung erreicht. Dabei sind fachlich und technisch hoch integrierte und komplexe Lösungen entstanden, die nicht länger als lose Folge von Einzelprojekten betrachtet werden können, sondern einer zentralen Regulation und Koordination bedürfen. Als Folge daraus wurden in den Unternehmen Governance-Strukturen für den BI-Bereich entwickelt, denen es jedoch bislang an einer ausreichenden konzeptionellen Unterstützung für die Erfassung, Überwachung und Steuerung der zugrunde liegenden BI-Landschaft fehlt. Anhand eines Grundkonzepts zur Service-basierten Abbildung von BI-Infrastrukturen werden im Rahmen dieses Beitrags zwei Fallstudien analysiert und diskutiert. Ergebnis des Beitrags ist die Darstellung eines service-basierten Konzepts zur Unterstützung von BI-Governance, das insbesondere die Aspekte der Komposition und des Managements des Lebenszyklus von Services berücksichtigt.

1 Relevanz und Motivation

Für eine zunehmende Zahl von Unternehmen hat Business Intelligence (BI) als leistungsfähige IT-basierte Managementunterstützung einen zentralen Stellenwert eingenommen [Ga06, Me04]. Wachsende Anforderungen an eine hochwertige und zeitnahe Informationsversorgung erfordern BI-Infrastrukturen zur Unterstützung sämtlicher Managementebenen eines Unternehmens, wobei mittlerweile zusätzlich eine konsequente Ausrichtung der Analysen auf transaktionsnahe Geschäftsprozessdaten angestrebt wird [Ec06a, Le06].

Aus diesen Entwicklungen ergeben sich erhebliche Herausforderungen an Erstellung und Betrieb geeigneter BI-Lösungen. Üblicherweise werden dazu integrierte BI-Lösungen auf Basis mehrstufiger Architekturen entwickelt, die auf heterogene Quellsysteme aufsetzen [Ba06, KMU06], große Datenvolumina beherrschen sollen [Wi05] und hohen Anforderungen an Datenintegration und Datenharmonisierung genügen müssen [En99].

Business Intelligence (BI) wurde ursprünglich als Marketing-orientierter Begriff von Beratern und Analysten eingeführt. Das Marktforschungsunternehmen Gartner, Inc. formulierte z.B. in den 1990er Jahren: "... *data analysis, reporting, and query tools can help business users wade through a sea of data to synthesize valuable information from it – today these tools collectively fall into a category called 'Business Intelligence'*" [ASA03]. In den Folgejahren heftig diskutiert, hat sich der Begriff inzwischen in Theorie und Praxis etabliert [z.B. CG04, Hi01, KM02, Me02] und wird heute allgemein als „integrierter, unternehmensspezifischer, IT-basierter Gesamtansatz zur betrieblichen Entscheidungsunterstützung“ angesehen [KMU06].

Inzwischen haben Lösungen zur betrieblichen Entscheidungsunterstützung grundlegende Veränderungen erfahren. In mehr und mehr Unternehmen basieren BI-Lösungen auf komplexen und vielschichtigen Architekturen mit zahlreichen Anbindungen an operative Transaktionssysteme [Mc07, Sh07]. Vormalig isolierte BI-Lösungen sind zu unternehmensweiten Infrastrukturen zusammengewachsen, die alle Hierarchieebenen und Funktionsbereiche [MA03] umgreifen. Vor dem Hintergrund dieses infrastrukturellen, integrierten und kostenintensiven Charakters aktueller BI-Ansätze werden neue organisatorische Anforderungen ersichtlich, die weit über die Koordination einer einfachen Folge von isolierten Entwicklungsprojekten hinausgehen. Dies bestätigen auch die Ergebnisse empirischer Untersuchungen, die zeigen, dass BI-Anwender die frühen Entwicklungs- und Konzeptionsphasen ihrer BI-Lösungen verlassen und eine Phase des qualitativen und quantitativen Wachstums sowie der Konsolidierung erreicht haben [UK08].

Diese Entwicklungen bringen die Notwendigkeit mit sich, individuelle Konzepte zu entwickeln, die die Institutionalisierung der BI-Bereiche und Aufgaben regeln sowie die Bereitstellung und den Betrieb von BI-Anwendungen gestalten [z.B. CE05, Ec06b, Ge07, MBG06, SH03, To06]. Diese Einschätzung wird zunehmend auch von der Praxis vertreten: Vielfach haben Unternehmen mit der Einrichtung von Organisationseinheiten zur Professionalisierung der BI begonnen. Diese Einheiten variieren jedoch stark in Größe, Zielsetzung und Ursprung – manche werden vorwiegend als Teil der IT wahrgenommen, während andere eng mit den Fachbereichen verbunden sind [UK08]. Aufgrund der relativen Neuartigkeit dieser Initiativen gibt es einen erkennbaren Mangel an Konzepten, die den BI-Einheiten helfen, die Komplexität der BI-Steuerung anzugehen. Es ist zu berücksichtigen, dass BI von einer vielschichtigen Wechselbeziehung fachlicher und technischer Anforderungen geprägt wird, wobei zahlreiche heterogene Fach- und Funktionsbereiche zusammengeführt werden müssen. Die Anforderungen gehen dabei weit über jene im Bereich traditioneller Transaktionssysteme hinaus. Zunehmende Komplexität entsteht dadurch, dass BI-Lösungen immer häufiger die Grenzen einzelner Organisationseinheiten und sogar von Unternehmen überschreiten.

Für die Entwicklung ganzheitlicher und nachhaltiger Konzepte kann der BI-Bereich von den Erfahrungen und Ergebnissen zur Theorie und Praxis genereller IT-Governance-Ansätze profitieren, die in der jüngeren Zeit zunehmend Aufmerksamkeit erfahren haben [z.B. Al06, BG05, DV06, Hi05, La00, Se03, WR04]:

IT-Governance umfasst die Gestaltung von Regelungen, Empfehlungen, Rollen und Verantwortungszuweisungen sowie Projektpriorisierungen, mit dem Ziel einer konsistenten Ausrichtung der IT an der Strategie des Unternehmens [Br05, Gu08, IGI03]. Dabei ist zu empfehlen, die genannten Besonderheiten und Herausforderungen von BI durch einen dedizierten BI-Governance-Ansatz zu adressieren. Erste Schritte in dieser Richtung sind die Entwicklung von BI-Reifegradmodellen sowie eine BI-spezifische Adaption von IT-Servicemanagement-Konzepten [Ec08, He07, Ph05]. Trotz dieser Bemühungen besteht immer noch ein Mangel an Konzepten, die eine Abgrenzung, Beschreibung, Messung und Steuerung von BI-Leistungen und BI-Services angemessen unterstützen.

Dieser Beitrag stellt anhand von zwei Fallstudien über umfangreiche und ausgereifte BI-Lösungen die Ergebnisse der Validierung eines Konzepts zur Erfassung und Abbildung von BI-Services dar. Daraus abgeleitet zeigt der Beitrag die erforderlichen Erweiterungen des Konzepts für das Management von BI-Services und BI-Lösungen auf, das zugleich als zentraler Baustein für den Entwurf eines umfassenden, werkzeuggestützten BI-Governance-Ansatzes verstanden wird.

2 Begriffe und bestehende Konzepte zur Governance

Um den Begriff BI-Governance definieren und einen geeigneten Ansatz zur Analyse der nachfolgenden Fallstudien entwickeln zu können, sind zunächst wesentliche Aspekte der übergeordneten IT-Governance herauszuarbeiten. Gemäß Weill ist IT-Governance zu verstehen als *“...specifying the framework for decision rights and accountabilities to encourage desirable behavior in the use of IT”* [We04]. Ähnlich definiert das IT Governance Institute den Begriff als *“an integral part of enterprise governance and consists of the leadership and organisational structures and processes that ensure that the organisation’s IT sustains and extends the organisation’s strategies and objectives”* [IGI03]. Neben dem Entwurf von Strategien und der Gestaltung einer strategisch ausgerichteten Aufbau- und Ablauforganisation, betrachtet Van Grembergen auch die Entwicklung von unterstützenden Kennzahlen als einen wichtigen Baustein der IT-Governance [Va03].

Hierbei sind die Begriffe „IT-Governance“ und „IT-Management“ zu unterscheiden. Gemäß Weill sind konkrete Steuerungsentscheidungen nicht der IT-Governance, sondern dem IT-Management zuzuordnen. IT-Governance legt vielmehr systematisch fest, wer, an welcher Stelle einen bestimmten Typ von Entscheidungen trifft (Entscheidungsrecht), wo und bei wem ein Recht zur Mitsprache liegt (Mitspracherecht), und wie die Wahrnehmung der damit verbundenen Verantwortung gemessen wird [We04]. Weill stellt außerdem fest, dass eine Identifikation der wesentlichen Entscheidungstypen maßgeblich für die Gestaltung eines IT- bzw. BI-Governance-Konzepts ist [We04]. Diese Einschätzung wird hier ausdrücklich geteilt. Mehr noch ist es als notwendige Voraussetzung anzusehen, zuerst die Objekte und Ressourcen zu erfassen, die diesen Entscheidungstypen zugrunde liegen bzw. auf die sie sich beziehen [BJZ92].

Die Entscheidungsobjekte der IT-Governance sind hierbei die einzelnen mittels IT bereitgestellten Services. Eine Identifikation dieser Services kann anhand von Konzepten des IT-Servicemanagements (ITSM) erfolgen, die gleichsam als Fundament für umfassende IT-Governance-Konzepte zu verstehen sind. Gemäß dem Office of Government Commerce (OGC) ist ein „IT-Service“ definiert als Mittel, dem Kunden einen Wertbeitrag zu leisten, indem die Erreichung des vom Kunden gewünschten Ziels unterstützt wird, ohne ihm die Verantwortung für Kosten und Risiken zu übertragen, die für die Erbringung des Wertbeitrags erforderlich sind [OGC07].

ITSM leistet eine Integration von Fachbereich und IT auf Ebene der Geschäftsprozesse und unterstützt so eine Konkretisierung der IT-Governance-Ansätze. Dies geschieht durch Bereitstellung eines Konzepts zur Dekomposition komplexer IT-Organisationen und IT-Lösungen in wohldefinierte und gekapselte Services. Diese sind eine entscheidende Grundlage für die Ableitung von adäquaten Leistungskennzahlen und fachlich orientierten Service-Level-Agreements (SLAs). Anhand der Kennzahlen kann wiederum eine wirkungsvolle Überwachung und Steuerung von internen und externen IT-Services erfolgen [IGI03].

Die Grundprinzipien der IT-Governance können aus der Corporate-Governance abgeleitet werden [We04]. Dementsprechend ergibt sich eine BI-Governance aus der anforderungsspezifische Interpretation und Ableitung der Prinzipien und Ziele der IT-Governance [Gu08]. Diese Beziehung verdeutlicht Abb. 1.

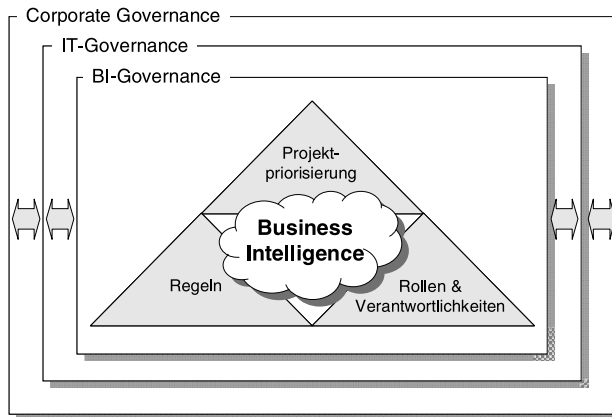


Abbildung 1: Hierarchische Ableitung der BI-Governance-Prinzipien
(in Anlehnung an [Gu08])

Zusammenfassend bezieht sich „BI-Governance“ auf die Sicherstellung und Unterstützung der Ausführung von Geschäftsprozessen sowie der Bereitstellung der dazu erforderlichen BI-Services, um dadurch eine zielgerichtete, entscheidungsunterstützende Informationsversorgung zu gewährleisten. Dies geschieht durch die direkte Erstellung der entscheidungsunterstützenden Informationen oder durch Entwurf, Betrieb bzw. Unterstützung dazu geeigneter BI-Systeme. Die Organisationseinheiten, die BI-Services bereitstellen, übernehmen volle Verantwortung für Kosten und Risiken bei Entwicklung und Betrieb der erforderlichen IT-Systeme.

BI-Service-Konzept

Ausgangspunkt dieser Arbeit ist ein Grundkonzept zur Differenzierung von BI-Services. Es wurde im Kontext der Forschung zu BI-Outsourcing entwickelt und durch eine empirische Studie zu diesem Thema motiviert. Die Ergebnisse dieser Studie zeigen auch das Nutzenpotential einer unternehmensübergreifenden Verteilung von BI-Services [BHK07].

Das Konzept spannt drei Dimensionen auf, die für die Differenzierung von BI-Services als erforderlich angesehen werden: Eine Komponenten-Dimension, eine Fachlichkeits-Dimension und eine Lebenszyklus-Dimension.

Ausgangspunkt für die Betrachtung von BI-Services sind die Software-Komponenten, die den Kern jeder BI-Lösung bilden: BI-Services adressieren oder verwenden in der Regel spezifische Subsysteme, wie z.B. Data-Warehouses, Data-Mining- oder OLAP-Anwendungen. Um diese unterschiedlichen Software-Komponenten innerhalb einer Komponenten-Dimension zu gliedern, kann auf etablierte BI-Architekturordnungsrahmen zurückgegriffen werden, vgl. hierzu z.B. [KB06]. Die Relevanz einer Komponentendimension ergibt sich aus der Unterschiedlichkeit und Abgeschlossenheit der Komponenten sowie aus den divergierenden Hardware- und Software-Anforderungen.

Die Dimension „Fachlichkeit“ unterscheidet infrastrukturorientierte und damit eher technische Services, wie z.B. die Bereitstellung von Rechnerkapazitäten oder Speicherlösungen, von eher fachlich orientierten Services, wie z.B. Harmonisierung von Stammdaten oder der Entwicklung einer individuellen Data-Mining-Lösung. Das Kernkriterium zur Unterscheidung von Services entlang dieser Dimension ist das Ausmaß an fachlicher Verantwortung, das ein Service-Nutzer einem Service-Anbieter überträgt. Je mehr Verantwortung für den eigentlichen Inhalt auf den Service-Anbieter übertragen wird, desto mehr muss dieser mit der fachlichen und kontextspezifischen Semantik des Anwenders vertraut sein. Aufbauend auf dem Konzept des Service-Stacks von Kern, Willcocks et al. [KWL02] wird dazu im Folgenden die Differenzierung von vier Schichten vorgeschlagen:

- *Hardware* – Bereitstellung und Betrieb der relevanten Rechen-, Speicher- und Telekommunikationsinfrastruktur, die notwendig ist, um eine oder mehrere BI-Komponenten zu bedienen.
- *Tools* – dies bezieht sich auf die eigentliche BI-Software, von ETL-Werkzeugen und Data-Warehouse-Applikationen bis hin zu spezialisierten Lösungen für Reporting und Visualisierung.
- *Templates* – verstanden als vorkonfigurierte Anwendungen und vorbereitete Inhalte, die an einzelne Bedarfe angepasst werden können.
- *Content* – diese Schicht betrifft die eigentliche fachliche Semantik. Ein Anbieter auf dieser Ebene verantwortet die Definition, Sammlung, Strukturierung, Transformation und/oder Darstellung der entscheidungsrelevanten Daten.

Die dritte erforderliche Dimension betrachtet die Phasen im Lebenszyklus der BI-Lösung: Es ist zu unterscheiden, ob ein Service primär auf die Entwicklung bzw. die Bereitstellung einer Komponente oder Teilen davon ausgerichtet ist oder auf deren Betrieb. Obwohl diese Dimension noch feiner gegliedert werden könnte, liefert die Unterscheidung von Entwicklung bzw. Bereitstellung und Betrieb einen vorrangigen Erkenntnisgewinn, da diese Phasen grundlegend verschiedene Werkzeuge, Fertigkeiten und Managementmethoden verlangen. Die Lebenszyklus-Dimension ist für alle Komponenten und alle Ebenen der Fachlichkeits-Dimension relevant.

Das resultierende BI-Service-Konzept ist in Abb. 2 dargestellt.

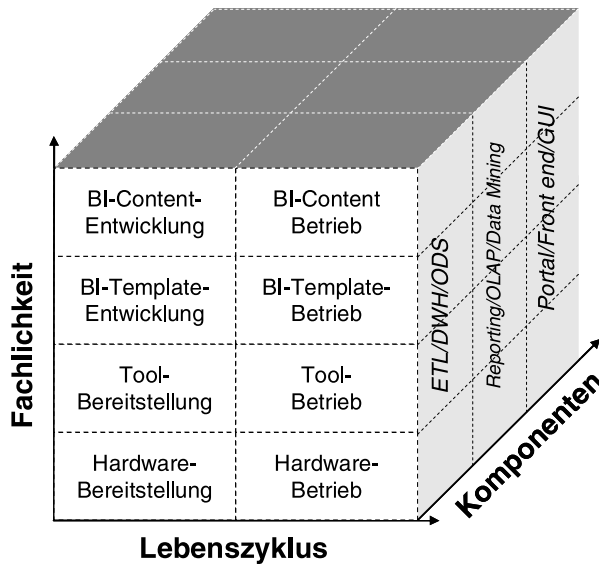


Abbildung 2: Konzept zur Differenzierung von BI-Services

Forschungsmethodik

Die nachfolgend diskutierte Untersuchung diente einer Validierung des BI-Service-Konzepts unter den Bedingungen eines realen Anwendungskontexts, der Bewertung von möglichen Anwendungsbereichen sowie der Identifikation von notwendigen Erweiterungen und Anpassungen. Aufgrund dieser explorativen Ausrichtung wurde eine qualitative Fallstudienanalyse als angemessene Forschungsmethodik gewählt.

Die Identifikation und Kontaktierung der Unternehmen basiert auf den Ergebnissen einer vorangegangenen quantitativen Studie. Im Hinblick auf die genannten Ziele wurden zwei Fälle anhand der Kriterien „Reifegrad des BI-Ansatzes“, sowie „Größe und Komplexität der BI-Lösung“ ausgewählt. Zugleich verfolgen die Unternehmen mit ihrer BI komplementäre Ansätze. In beiden Fällen handelt es sich um global agierende Unternehmen in wettbewerbsintensiven Branchen.

Die Erhebung der Datengrundlage für die Fallstudien erfolgte durch Interviews, Workshops sowie die Analyse verfügbarer und einschlägiger Dokumente. Die angeführten Methoden wurden im Hinblick auf die genannten Ziele und das dargestellte BI-Service-Konzept ausgewählt.

2.1 Fallstudie 1: Strukturierung und Integration heterogener BI-Lösungen bei einem Automobilhersteller

Der erste Fall thematisiert die Herausforderungen für das Management der BI-Infrastruktur eines Automobilherstellers. Das betrachtete Unternehmen ist mit einer Vielfalt heterogener BI-Systeme in den verschiedenen Geschäftsbereichen konfrontiert. Die Systeme sind weitgehend unabhängig voneinander konzipiert und in einer Reihung von Projekten ohne weitergehende projektübergreifende Abstimmungen entwickelt worden. Um eine effiziente Verzahnung mit der vorhandenen ERP-Systemlandschaft zu ermöglichen ist die Data-Warehouse-Plattform bei allen Systemen identisch (SAP BW). Dennoch laufen die diversen Lösungen zum Teil in verschiedenen Instanzen, besitzen individuelle Anpassungen und werden durch verschiedene Reporting-Lösungen ergänzt.

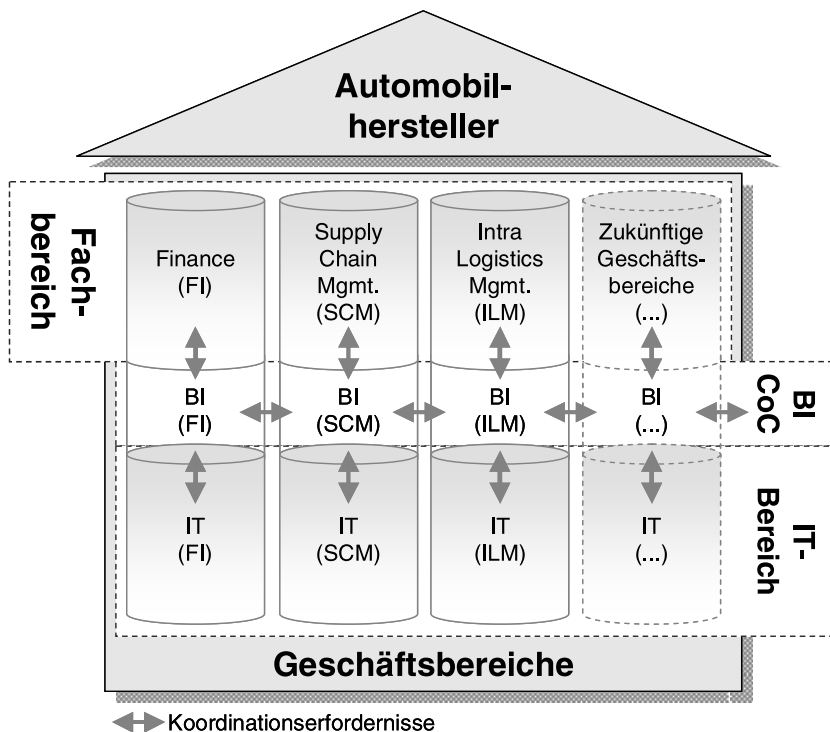


Abbildung 3: BI-bezogener Koordinationsbedarf des Automobilunternehmens

Bisher ist jedes Projekt in Kooperation zwischen dem jeweiligen Geschäftsbereich und der zentralen IT-Abteilung durchgeführt worden. Im Laufe der Zeit nahmen Anzahl,

Größe und Bedeutung der BI-Lösungen immer mehr zu. Dabei zeigten sich Abhängigkeiten zwischen den verschiedenen Systemen, die zugleich Ursache für erkennbare Redundanzen waren. Die Entscheidungsträger im Unternehmen schätzten den Ansatz entsprechend als zunehmend ineffizient und ineffektiv ein. Zugleich konnte die allgegenwärtige und etablierte Ad-hoc-Koordination zwischen den Interessengruppen nicht länger aufrecht erhalten werden.

Als Antwort auf diese Situation wurde im Unternehmen ein sogenanntes BI CoC (Center of Competence) innerhalb der IT-Abteilung etabliert, das als Verbindungsglied zwischen den verschiedenen Geschäftsbereichen und der zentralen IT agiert (vgl. Abb. 3).

In einer ersten Analyse identifizierte das BI CoC eine Anzahl von zentralen Herausforderungen, die im Folgenden auf Grundlage des oben eingeführten Konzepts strukturiert und erörtert werden.

Hardware-Bereitstellung und -Betrieb

Hardwareseitig leiden die Lösungen der Geschäftsbereiche unter dem intransparenten und ungeplanten Wachstum – sowohl Größe und Datenvolumen als auch den Umfang der Nutzung betreffend. Mittlerweile ist das Antwortverhalten bei mehreren Lösungen inakzeptabel.

Der BI-Governance-Aspekt dieses Problems liegt in der Notwendigkeit, Strukturen und Prozesse einzuführen, die Entwurf, Überwachung und Steuerung der Hardware-Services sicherstellen, die für das gesamte Spektrum an BI-Lösungen erforderlich sind – jeweils unter Berücksichtigung einer effizienten Ressourcenallokation. Dies bedeutet wiederum, dass Informationen erforderlich sind, die notwendige Hardware-Bereitstellungs-Services spezifizieren und Anforderungen, der fachlichen Ebenen konkretisieren.

Tool- bzw. Templatebereitstellung und -betrieb

Die Bereitstellung und Entwicklung der fachlichen Anwendungen ist geprägt durch eine parallele und teilweise redundante Realisierung von Reporting-Komponenten mit gleichartiger Funktionalität (z.B. Vorlagen, Oberflächen-Designs). Darüber hinaus gibt es teilweise auch eine fachlich nicht erforderliche Heterogenität in der Ausgestaltung der Lösungen. Das alles führt zu vermeidbaren Effizienzverlusten und verlangt nach einer eingehenden Untersuchung der Möglichkeit, standardisierte Vorlagen und Anwendungen zu definieren, die wiederverwendet oder parallel genutzt werden können. Eine Voraussetzung dafür ist die Entkopplung des vielfach hochindividuellen Inhalts von der eigentlichen Reporting-Software.

Contententwicklung und -betrieb

Eine Standardisierung ist darüber hinaus auch im Hinblick auf die Inhalte erforderlich: Unterschiedliche und unzureichend harmonisierte Vorgehensmodelle für die Entwicklung von BI-Lösungen verursachen inkonsistente Datenarchitekturen und heterogene Namenskonventionen. Dies zieht vermeidbare syntaktische und semantische Inkompatibilitäten und inkonsistente Berichte nach sich und versperrt so den Weg für geschäftsbe-

reichsübergreifende Analysen. Ein weiterer Problembereich sind fehlende Informationen zum Eigentümer von und zur Verantwortung für Quelldaten.

Auch hier zeigt sich somit ein Mangel an Strukturen und Prozessen durch die definierte Bausteine für BI-Lösungen vorgegeben werden. Auf diesen Bausteinen beruhen standardisierte und eindeutig spezifizierte Vorlagen und Anwendungen, die über verschiedene Einzellösungen hinweg genutzt werden können. Relevante Information beziehen sich dabei auf die bereitzustellenden Daten sowie die zugehörigen Datenmodelle und Metadaten, insbesondere zu Volumen, Semantik, Eigentümer und Verwendung der Daten.

Lebenszyklus – Betriebsphase

Der derzeitige Betrieb wird durch eine unzureichende Dokumentation bezüglich der Verantwortung für System-Administration und -Koordination erschwert. Weiterhin stellen die aus der zeitzoneübergreifenden und intensiven Nutzung der BI-Lösungen resultierenden kurzen Zeitfenster für technische Wartung und Aktualisierung, eine enorme Herausforderung dar.

Diskussion

Im Fall wurden die strategische Relevanz und unternehmensweite Bedeutung von bereichsübergreifend integrierten BI-Systemen nicht vorausgesehen. Mit dem Wachstum der einzelnen Lösungen einher geht auch ein zunehmender Bedarf an adäquaten BI-Governance-Regularien. Es muss jedoch darauf hingewiesen werden, dass die erörterten Probleme nicht mithilfe einer Liste an einfachen Vorschriften behoben werden können. Selbst die Vertreter des BI CoC betonen, dass die Unterschiede zwischen den vorhandenen BI-Lösungen nur teilweise in einer fehlenden Regulierung begründet sind. Der überwiegende Teil der Unterschiede kann mit den individuellen Anforderungen der Geschäftsbereiche gerechtfertigt werden. Die resultierenden Herausforderungen beschreibt der Leiter des BI CoC folgendermaßen:

„Was muss reguliert werden, und welche Bestimmungen sind notwendig und angemessen? Und wie sagen wir unseren alten Kunden, dass sie nun nicht mehr das tun dürfen, was sie bislang getan haben?“

Der Fall verlangt nach einem modularen Ansatz, der darauf abzielt, wiederverwendbare und standardisierte Services zu identifizieren und isolieren. Er zeigt, dass es Bedarf für eine koordinierende Einheit gibt, die einen Überblick über das Spektrum an BI-Services behält und die deren Zusammenstellung zu integrierten Lösungen unter Berücksichtigung der Wiederverwendung bestehender Services sicherstellt. Das dargestellte Konzept zur Differenzierung von BI-Services (vgl. Abb. 2) wird dabei als geeigneter Ausgangspunkt für die Bewältigung der übergeordneten *Kompositionsaufgaben* verstanden, ist jedoch in dieser Hinsicht zu erweitern.

2.2 Fallstudie 2: Dokumentation und Lebenszyklus-orientierte Unterstützung der BI-Lösung einer Börsenorganisation

Der zweite Fall hebt vor allem Lebenszyklusaspekte von BI-Lösungen hervor. Bei dem betrachteten Unternehmen handelt es sich um eine ehemals nationale Börse, die sich im Laufe der Zeit zu einer der größten Börsenorganisationen der Welt entwickelt hat. Ende der neunziger Jahre sah sich die Börse zum einen mit Veränderungen der Unternehmensumwelt konfrontiert, die vor allem auf das Wachstum und die hohe Dynamik der Finanzmärkte im Rahmen des internetbasierten Handels zurückzuführen waren, und zum anderen mit zunehmenden Forderungen interner und externer Interessengruppen nach präziseren, zuverlässigeren und aktuelleren entscheidungsunterstützenden Informationen über Finanzmärkte. Die hierfür entwickelte BI-Lösung basiert auf einem Real-time-Data-Warehouse, das als zentrale Drehscheibe für die Informationsdistribution fungiert. Insbesondere erlaubt die Lösung internen und externen Nutzern eine zeitnahe Analyse finanzmarktbezogener Massendaten.

Die Dynamik auf den weltweiten Finanzmärkten spiegelt sich in der Komplexität der Lösung wider. Hierbei ist eine hohe Flexibilität bei der Bereitstellung entscheidungsrelevanter Informationen von immenser Bedeutung. Für interne wie externe Kunden ist die Länge der Zeitdauer zwischen Identifizierung eines informationellen Defizits und Bereitstellung der erforderlichen Informationen ein kritischer Erfolgsfaktor. Der nachhaltige Erfolg der Börseorganisation hängt im Hinblick auf BI von der Fähigkeit ab, die komplexen Prozesse der Anforderungsanalyse, der Umsetzung und des Betriebs von kundenindividuellen BI-Lösungen professionell zu managen.

Um diesen anspruchsvollen Anforderungen zu genügen wurde die BI-Lösung als Best-of-Breed-Ansatz realisiert und aus den Produkten verschiedener Anbieter zusammengestellt (Informatica, Microstrategy, Microsoft etc.). Betrieb und Wartung der Lösung benötigen daher unterschiedlichste Qualifikationen und eine intensive Koordination der beteiligten Mitarbeiter. Zur Sicherstellung höchster Professionalität und Flexibilität bündelte das Unternehmen BI-Wissen und BI-Verantwortung in einem dedizierten Center of Competence (CoC). Diese Einheit übernimmt alle Aufgaben bezüglich der Real-time-BI-Lösung. Lediglich die Hardware-Infrastruktur wird von einem unternehmensinternen Rechenzentrum bereitgestellt.

Mit der Zeit wurden die Anforderungen an die Bereitstellung von Informationsprodukten immer heterogener und anspruchsvoller. Zusätzliche Herausforderungen an die Fähigkeit des CoC zur schnellen Konzipierung, Entwicklung und Umsetzung neuer Lösungen, ergaben sich aus der vorgesehenen Umgestaltung des CoC in einen am externen Markt agierenden Outsourcing-Anbieter von BI-Services. Entsprechend dem vorgestellten Konzept ergeben sich die folgenden Schlüsse:

Die Services zur Hardware-Bereitstellung werden durch das interne Rechenzentrum verantwortet und über fest vereinbarte Service-Levels gesteuert. Im vorliegenden Fall sind die zwei untersten Ebenen der Dimension „Fachlichkeit“ (vgl. Abb. 2) entsprechend über gut definierte Schnittstellen entkoppelt.

Bis vor Kurzem lagen die Services des CoCs überwiegend auf der Ebene der Bereitstellung und dem Betrieb von *Content* – z.B. die Entwicklung von Informationsprodukten wie Marktindizes oder OLAP-Cubes für finanzielle Analysen. Die Börsenorganisation war primär für Korrektheit und rechtzeitige Bereitstellung der Informationsprodukte verantwortlich.

Es stellte sich jedoch heraus, dass die Ebene der Bereitstellung von Templates und Tools nicht im gleichen Maß definiert war, da die zugehörigen Aktivitäten bislang effektiv durch das eingespielte Team des CoC übernommen wurden und keine weitergehende formale Koordination erforderlich war. Weiterer Koordinationsbedarf wurde zudem durch die Verwendung eines in sich stimmigen Pakets von optimierten Softwarelösungen und einer etablierten Verteilung der Aufgaben und Spezialisierungen innerhalb des Teams abgefangen.

Mit der gestiegenen Vielfalt an Informationsprodukten und der Anforderung als Anbieter für Entwurf, Umsetzung und Betrieb von BI Lösungen für externe Kunden aufzutreten, wurde der Mangel an Transparenz auf der Anwendungs-Ebene zu einem massiven Hindernis. Neue Services mussten technisch und organisatorisch in die bestehende BI-Umgebung integriert und im Hinblick auf die gegebene Infrastruktur gestaltet werden.

Diskussion

Anders als im ersten Fall lagen die Herausforderungen für die Börsenorganisation *nicht* in der Zusammenstellung von Lösungen auf Grundlage einer heterogenen Systemlandschaft sondern in der Dekomposition einer ursprünglich homogenen Lösung, um auf heterogene Anforderungen innerhalb kurzer Zeit reagieren zu können.

Die Börsenorganisation hat das hier diskutierte Konzept bereits umgesetzt, um damit die vorhandene BI-Service-Landschaft zu strukturieren. Bei der Entwicklung neuer Lösungen anhand des Konzepts wurde jedoch offensichtlich, dass weiterer Entwicklungsbedarf besteht. Der Leiter des CoCs erklärte:

„Wir mussten feststellen, dass wir die Service-Beschreibungen nicht von Anfang an vollständig ausfüllen können, weil die Kunden, mit denen wir es zu tun haben, normalerweise ihren Bedarf vorher nicht genau kennen oder ihnen die notwendigen Vorkenntnisse fehlen, um die Anforderungen an solche komplexe BI-Lösungen ausreichend zu definieren.“

Die CoC-Mitarbeiter betonten, dass es einen erkennbaren Bedarf für eine service-basierte Dokumentation gibt, die es ermöglicht, die jeweils aktuelle Zusammenstellung und Konfiguration der einzelnen BI-Lösungen aus Kunden- und Service-Level-Perspektive zu beschreiben. Zugleich forderten sie eine IT-basierte Unterstützung ein, um die Services auf allen Ebenen und hinsichtlich aller Abhängigkeiten erfassen und verfolgen zu können.

2.3 Vergleich der Fallstudien

Die Hauptunterschiede zwischen beiden Fallstudien liegen in der Struktur der gegenwärtigen BI-Lösungen, ihrer jeweiligen Historie und den darin begründeten gegenwärtigen Problemen.

In beiden Fälle führte eine gesteigerte Komplexität zu einer Nachfrage nach einer systematischen Darstellung von Services, die als modulare Bausteine für individuelle Lösungen für eine Vielzahl von internen oder externen Kunden dienen können. Weiterhin werden beide Fällen durch die Interaktion verschiedener Partner geprägt. Im Falle des Automobilunternehmens sind dies das BI CoC, die IT-Einheiten sowie die unterschiedlichen Geschäftsbereiche. Bei der Börsenorganisation handelt es sich um das CoC, das Rechenzentrum, die internen und externen Nutzer der Informationsprodukte sowie die Nutzer der Outsourcing-Services.

Die Fallstudien unterstreichen die Annahme, dass eine Kernanforderung für BI-Governance und das Management von BI-Lösungen in der Erlangung eines vollständigen und transparenten Überblicks über die erforderlichen Services und deren Zusammenspiel liegt. Im Fall des Automobilherstellers zeigt sich darüber hinaus die Notwendigkeit, die Komposition der Services zu komplexen Lösungen unter besonderer Berücksichtigung der Wiederverwendung von Services zu betrachten. Der Fall der Börsenorganisation zeigt, dass eine Konzeption für das Management von BI-Services auch die schrittweise Sammlung von Informationen unterstützen muss, angefangen bei den ersten Phasen der Anforderungsanalyse bis zur Verfolgung dieser Anforderungen über den Lebenszyklus.

Die Unterschiede zwischen den Fällen sind in Tabelle 3 gegenübergestellt.

	Fallstudie 1: Automobilunternehmen	Fallstudie 2: Börsenorganisation
Mit dem CoC interagierende Partner	<ul style="list-style-type: none"> ▪ Geschäftsbereiche (Nutzer) ▪ IT-Einheiten 	<ul style="list-style-type: none"> ▪ Rechenzentrum ▪ Interne Kunden der Informationsprodukte ▪ Externe Kunden der Informationsprodukte ▪ Kunden von ASP-Lösungen
Fokus	<ul style="list-style-type: none"> ▪ Zusammenstellung einer Lösung basierend auf modularen Services zur Steigerung der Effizienz 	<ul style="list-style-type: none"> ▪ Dekomposition der Lösung zur Erhöhung der Flexibilität ▪

Herausforderungen	<ul style="list-style-type: none"> ▪ Erlangung von Transparenz bezüglich Tools, Content und Hardware-Anforderungen ▪ Wiederverwendung von Hardware ▪ Wiederverwendung von Reporting-Tools und -Templates ▪ Integration von Datenmodellen ▪ Informationen zu Rollen und Verantwortlichkeiten 	<ul style="list-style-type: none"> ▪ Erlangung von Transparenz hinsichtlich der Tool-Bereitstellung ▪ Verkürzung der Zeit zur Erfassung von Kundenanforderungen und zur Entwicklung neuer Lösungen ▪ Definition in sich geschlossener Services zur Bereitstellung von Tools
Kritische Ebenen der Fachlichkeits-Dimension	<ul style="list-style-type: none"> ▪ Hardware ▪ Tools/Templates (insb. Reporting-Komponenten) ▪ Content 	<ul style="list-style-type: none"> ▪ Tools/Templates
Anforderungen an das Konzept	<ul style="list-style-type: none"> ▪ Unterstützung der Zusammenstellung von Lösungen 	<ul style="list-style-type: none"> ▪ Unterstützung der Erfassung von Informationen hinsichtlich der Anforderungen ▪ Bereitstellung von IT-basierten Werkzeugen

Tabelle 3: Vergleich der beiden Fallstudien

3 Ableitung eines Konzepts für das BI-Service-Management

Auf Grundlage der obengenannten Anforderungen wird das Konzept zur Abbildung von BI-Services nachfolgend mit Konzepten für die Komposition von Services und für das Lebenszyklus-Management angereichert.

3.1 Komposition von Services: Zusammenhang von Solution und Service

Wie die Fälle zeigen, machen variierende Kundenanforderungen einen (De-)Kompositionsansatz für BI-Lösungen erforderlich. Lösungen für die Entscheidungsunterstützung sind hochgradig von den Anforderungen der Nutzer abhängig. Wie insbesondere durch den Fall des Automobilherstellers deutlich wird, können sich grund-

legende Anforderungen sogar innerhalb eines Unternehmens so stark unterscheiden, dass eine monolithische Lösung keine angemessene Option darstellt.

In erster Linie ist jedoch nicht der einzelne Service Gegenstand der Diskussion zwischen BI-Nutzern und BI-Service-Anbietern, sondern die *ganzheitliche BI-Lösung*. Aus Sicht der BI-Governance sind solche Lösungen jedoch als Bündel einzelner BI-Services zu verstehen, die zu einer umfassenden Lösung zusammengestellt werden müssen – unter Berücksichtigung von Konsistenz und Wiederverwendung. Nachfolgend wird diese Komposition von BI-Services als BI-Solution bezeichnet.

Eine solche BI-Solution könnte z.B. sein: „Eine *OLAP-Anwendung für die Marketingabteilung*“. Die BI-Solution umfasst BI-Services wie z.B. „*Bereitstellung der Datenbank-Hardware und -speicher-Infrastruktur für das Data-Warehouse*“, „*Wartung der OLAP-Anwendung*“ oder „*Datenextraktion, -transformation und Qualitätssicherung für die Bereitstellung von Vertriebsdaten*“. Die BI-Services dienen als Bausteine zur effizienten Gestaltung von BI-Solutions für eine Vielzahl von Kunden sowie zur Definition und Kommunikation von Rollen, Verantwortlichkeiten und Leistungskennzahlen.

Der Zusammenhang von Services und Lösungen ist in Abb. 4 dargestellt.

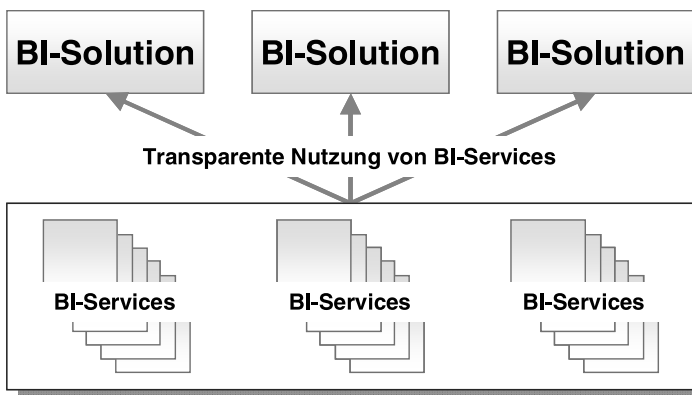


Abbildung 4: BI-Services und BI-Lösungen

Um die Komplexität dieses kompositionellen Ansatzes zu beherrschen, muss das BI CoC unterstützende Information über vorhandene und geplante Services erhalten – eine BI-Servicebibliothek (BI-Service-Library). Weiterhin sind Information erforderlich, die dokumentieren, welchen Benutzern welche Lösungen bereitgestellt wird und welche Service-Levels damit verbunden sind, sowie welche Rollen und Verantwortungen diesen zuzuordnen sind und welche Interdependenzen zwischen den Services bestehen.

3.2 Unterstützung und Verfolgung von BI-Services über deren Lebenszyklus

Die zweite zentrale Schlussfolgerung aus den Fällen ist, dass eine Notwendigkeit für die Unterstützung einer schnellen Informationsversorgung bezüglich der erforderlichen BI-Lösungen und der zugehörigen BI-Services besteht. Dazu ist erforderlich, eine Möglich-

keit zu schaffen, um schrittweise Services definieren zu können – ausgehend von einer eher umfassenden Lösungssicht, die nachfolgend schrittweise in individuelle und wohl-definierte Services heruntergebrochen wird. Die Services sollten dabei – wann immer möglich – aus der bestehenden Service-Bibliothek übernommen werden.

Um die Service-Bibliothek und die darin enthaltenen Informationen aktuell zu halten, müssen Änderungen kontinuierlich verfolgt werden. Wie im Bereich des IT-Service-Managements für operative Systeme, ist auch hier ein systematisches Change-Management zu implementieren, das die Prozesse und Strukturen für den Umgang mit Change-Requests definiert und gleichzeitig die Dokumentation der Services und Solutions auf dem aktuellen Stand hält.

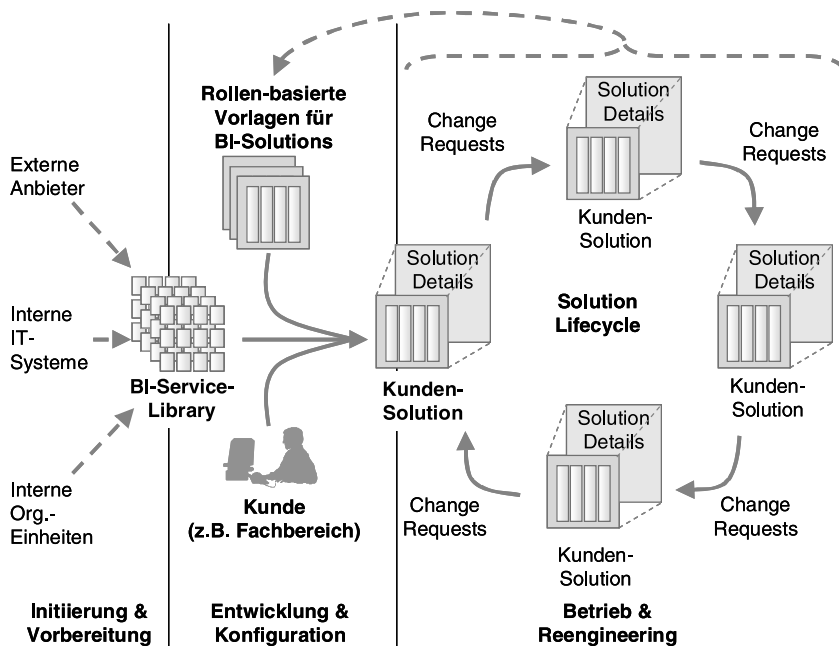


Abbildung 5: Resultierendes BI-Service-Management-Konzept

Der resultierende Ansatz ist in Abb. 5 dargestellt. Ausgangspunkt ist die *Initiierung & Vorbereitung* des BI-Service-Managements, bei dem bestehende BI-Systeme anhand des service-basierten Ansatzes erfasst, strukturiert und in der *BI-Service-Library* dokumentiert werden. Im weiteren Verlauf steht die Betrachtung der individuellen Anforderung des *Kunden* im Mittelpunkt.

Anhand *rollen-basierter Vorlagen für BI-Solutions*, die typische Anwendungsszenarios für BI beschreiben, wird eine Grundkonfiguration der *Kunden-Solution* ausgewählt und im Rahmen der *Entwicklung & Konfiguration* den Kundenanforderungen angepasst. Dabei erfolgt in erster Linie die Einbindung vorhandener Services aus der *BI-Service-Library*, in zweiter Linie die Anpassung bestehender Services und in dritter Linie die Aufnahme neuer Services.

Mit dem Ende von Entwicklung & Konfiguration stellt die Kunden-Solution mit den zugehörigen Service-Beschreibungen in Form der *Solution-Details* die umfassende Dokumentation der kundenindividuellen BI-Solution dar. Die Dokumentation wird über den *Solution-Lifecycle* angepasst und fortgeführt und bildet damit die Änderungen der Kunden-Solution aufgrund von *Change Requests* entsprechend ab. Ergeben sich aus den Kundenanforderungen grundsätzlich neue Zusammenstellungen von BI-Services, gehen diese letztlich in neue rollen-basierte Vorlagen für BI-Solutions über.

4 Ausblick

Das beschriebene Konzept wurde bereits mit Vertretern der Praxis diskutiert und wird derzeit im Rahmen weiterer Forschung konkretisiert. Basierend auf der Idee einer strukturierten Bibliothek von BI-Services und BI-Solutions liegt ein Hauptaugenmerk auf der Entwicklung eines Werkzeugs zur Unterstützung der

- Dokumentation bestehender Services und Solutions,
- Entwicklung von geplanten bzw. Reengineering bestehender Solutions,
- Kommunikation zwischen kooperierenden BI-Service-Anbietern und BI-Nutzern,
- Überwachung des aktuellen Zustands der Service-Landschaft und
- Analyse der Leistungsfähigkeit der BI-Einheiten.

Dabei soll ein BI-Werkzeug angewendet und eine OLAP-Lösung entwickelt werden, die eine flexible, multidimensionale Sicht auf diese Informationen zu Services, deren Verhalten und Zustand erlaubt.

Ein zweiter Forschungsbereich betrachtet die Integration des BI-Service-Management-Konzepts mit Ansätzen des IT-Service-Managements und mit den Konsequenzen für den Entwurf eines weiterreichenden BI-Governance-Ansatzes.

5 Literaturverzeichnis

- [Al06] Ali, S.: Effective Information Technology Governance Mechanisms: An Australian Study. In Gadjah Mada International Journal of Business, Heft 8, 2006; S. 69-102.
- [ASA03] Anandarajan, A. E. D. T.; Srinivasan, C. A.; Anandarajan, M. E. D. T.: Business Intelligence Techniques. Springer, Berlin, New York, 2003.
- [Ba06] Bange, C.: Werkzeuge für das Business Intelligence. In HMD – Praxis der Wirtschaftsinformatik, Heft 247, 2006, S. 63-73.
- [BJZ92] Boynton, A. C.; Jacobs, G. C.; Zmud, R. W.: Whose Responsibility Is IT Management?. In Sloan Management Review, Heft 33, 1992; S. 32-38.
- [BG05] Brown, A. E.; Grant, G. G.: Framing the Frameworks: A Review of IT Governance Research. In Communications of AIS, Heft 15, 2005; S. 696-712.

- [BHK07] Baars, H.; Horakh, T. A.; Kemper, H.-G.: Business Intelligence Outsourcing – A Framework. In (Österle, H.; Schelp, J.; Winter, R. Hrsg.): Proceedings of the 15th European Conference on Information Systems (ECIS 2007), St. Gallen, 2007; S. 1155-1166.
- [Br05] Broadbent, M.: Why Governance Matters. In CIO Insight, Heft 60, S. 29-30.
- [CE05] Cunningham, D.; Elliott, T.: The Burden of Trusted Information. In DM Review, Jg. 15, Heft 6; 2005, S. 18-33.
- [CG04] Chamoni, P.; Gluchowski, P.: Integrationstrends bei Business-Intelligence-Systemen: Empirische Untersuchung auf Basis des Business Intelligence Maturity Model. In Wirtschaftsinformatik, Jg. 46, Heft 2; 2004, S. 119-128.
- [DV06] De Haes, S.; Van Grembergen, W.: Information Technology Governance Best Practices in Belgian Organisations. In (IEEE Computer Society Hrsg.): Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS39), Kauai, HI, 2006, 195b.
- [Ec06a] Eckerson, W. W.: Performance Dashboards. Wiley, Hoboken, NJ, 2006.
- [Ec06b] Eckerson, W.: New Ways to Organize the BI Team. In Business Intelligence Journal Jg. 11, Heft 1, 2006, S. 43-48.
- [Ec08] Eckerson, W.: Gauge Your Data Warehousing Maturity. In <http://www.tdwi.org/publications/display.aspx?ID=7199>, Zugriff am 14.03.2008.
- [En99] English, L. P.: Improving Data Warehouse and Business Information Quality. Wiley, Hoboken, NJ, 1999.
- [Ga06] Gartner Inc.: Gartner Survey of 1,400 CIOs Shows Transformation of IT Organisation is Accelerating. In http://www.gartner.com/press_releases/asset_143678_11.html, Zugriff am 14.03.2008.
- [Ge07] Geiger, J. G.; Hill, B.; Loftis, L.; Ton, J. S.: Creating a BI Center of Excellence. In Business Intelligence Journal, Jg. 12, Heft 1, 2007.
- [Gu08] Gutierrez, N.: Business Intelligence (BI) Governance. In <http://www.businessintelligence.com/ex/asp/code.170/xe/article.htm>, Zugriff am 14.03.2008.
- [He07] Henschen, D.: HP Touts Neoview Win, Banking Solution, BI Maturity Model. In Intelligent Enterprise, Heft 10, 2007, S. 9.
- [Hi01] Hildebrand, K. (Hrsg.): Business Intelligence. HMD – Praxis der Wirtschaftsinformatik, Jg. 38, Heft 222, 2001.
- [Hi05] Hildreth, S.: Business in the Driver's Seat. In Computerworld, Heft 39, 2005; S. 31-33.
- [IGI03] IT Governance Institute: Board Briefing on IT Governance. Rolling Meadows, IL 2003.
- [KB06] Kemper, H.-G.; Baars, H.: Business Intelligence und Competitive Intelligence – IT-basierte Managementunterstützung und markt-/wettbewerbsorientierte Anwendungen. In HMD – Praxis der Wirtschaftsinformatik, Heft 247, 2006, S. 7-20.
- [KM02] Kemper, H.-G.; Mayer, R.: Business Intelligence in der Praxis – Erfolgreiche Lösungen für Controlling, Vertrieb und Marketing. Lemmens Medien, Bonn, 2002.
- [KMU06] Kemper, H.-G.; Mehanna, W.; Unger, C.: Business Intelligence – Grundlagen und praktische Anwendungen – Eine Einführung in die IT-basierte Managementunterstützung. 2. Aufl., Vieweg, Wiesbaden, 2006.
- [KWL02] Kern, T.; Willcocks, L. P.; Lacity, M. C.: Application Service Provision: Risk Assessment and Mitigation. In MIS Quarterly Executive, Heft 1, 2002, S. 113-126.

- [La00] Lainhart IV, J. W.: Why IT Governance Is a Top Management Issue. In *Journal of Corporate Accounting & Finance*, Heft 11, 2000, S. 33-40.
- [Le06] Lehmann, P.: Business Intelligence – Status Quo und zukünftige Entwicklungen. In *HMD – Praxis der Wirtschaftsinformatik*, Heft 247, 2006, S. 21-32.
- [MA03] Moss, L. T.; Atre, S.: *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. Addison-Wesley, Boston, MA, 2003.
- [MBG06] Miller, G. J.; Bräutigam, D.; Gerlach, S. V.: *Business Intelligence Competency Centers: A Team Approach to Maximizing Competitive Advantage*. Wiley, Hoboken, NJ, 2006.
- [Mc07] McKnight, W.: Moving Business Intelligence to the Operational World, Part 1. In *DM Review*, Heft 17, 2007; S. 28-28.
- [Me02] Mertens, P.: *Business Intelligence – ein Überblick*. Arbeitspapier an der Universität Erlangen-Nürnberg 2/2002, Nürnberg, 2002.
- [Me04] META Group Deutschland GmbH: *Business Intelligence Marktanalyse und Markttrends (Deutschland 2004) – Studienzusammenfassung für Cognos*. In http://www.cognos.com/at/news/pdf/metagroup_studienzusammenfassung.pdf, Zugriff am 14.03.2008.
- [OGC07] Office of Government Commerce: *Service Transition*. TSO (The Stationery Office), Norwich, 2007.
- [Ph05] Philippi, J.: Outsourcing und Offshoring von Business Intelligence-Lösungen. In (Schelp, J.; Winter, R. Hrsg.): *Auf dem Weg zur Integration Factory – Proceedings der DW2004 – Data Warehousing und EAI*. Physica, Heidelberg, 2005, S. 73-104.
- [Se03] Seabrook, D.: Cost Pressures Focus CIOs on Business. In http://www.gartner.com/2_events/symposium/2003/asset_53562_1115.jsp, Zugriff am 14.03.2008.
- [SH03] Strange, K. H.; Hostmann, B.: *BI Competency Center Is Core to BI Success*. Gartner, Inc., Stamford, 2003.
- [Sh07] Sherman, R.: Data Integration Trends: Future Shock. In *DM Review*, Heft 17. 2007, S. 23-23.
- [To06] Totok, A.: Entwicklung einer Business-Intelligence-Strategie. In (Chamoni, P.; Gluchowski, P. Hrsg.): *Analytische Informationssysteme. Business Intelligence-Technologien und –Anwendungen*, 3. Aufl.. Springer, Berlin, New York, 2006; S. 51-70.
- [UK08] Unger, C.; Kemper, H.-G.: Organisatorische Rahmenbedingungen der Entwicklung und des Betriebs von Business Intelligence – Ergebnisse einer empirischen Studie. In (Bichler, M.; Hess, T.; Krcmar, H.; Lechner, U.; Matthes, F.; Picot, A.; Speitkamp, B.; Wolf, P. Hrsg.): *Proceedings der Multikonferenz Wirtschaftsinformatik 2008 (MKWI2008)*, München, 2008.
- [Va03] Van Grembergen, W.: Introduction to the Minitrack IT Governance and its Mechanisms. In (IEEE Computer Society Hrsg.): *Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS36)*, Big Island, HI, 2003.
- [We04] Weill, P.: Don't Just Lead, Govern: How Top-Performing Firms Govern IT. In *MIS Quarterly Executive*, Heft 3, 2004, S. 1-17.
- [Wi05] Wiel, M. V. d.: Managing Large-scale Business Intelligence Solutions. In *Business Intelligence Journal*, Heft 10, 2005, S. 28-34.
- [WR04] Weill, P.; Ross, J. W.: *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*, Harvard Business School Press, Boston, 2004.

Business-Intelligence-Konzept auf Basis einer Event-Driven Service-Oriented Architecture

Tim Vogt, Sebastian Neuhaus, Markus Linden, Peter Chamoni

Fachbereich Betriebswirtschaft – Mercator School of Management

Lehrstuhl für Wirtschaftsinformatik und Operations Research

Universität Duisburg-Essen, Campus Duisburg

Lotharstraße 65, 47057 Duisburg

{tim.vogt | sebastian.neuhaus | markus.linden | peter.chamoni}@uni-due.de

Abstract: In Zeiten eines dynamischen Wettbewerbs ist eine schnelle Entscheidungsfindung erforderlich, die im Rahmen klassischer Data-Warehouse(DWH)-Systeme, als technische Umsetzung von Business Intelligence (BI), nur eingeschränkt ermöglicht wird [FB06, 54]. DWH-Ansätze, wie das Real-Time Data Warehousing, sollen den Anforderungen von Echtzeitunternehmen gerecht werden. Das Real-Time Data Warehousing impliziert eine schnellere Informationsversorgung als der klassische Ansatz. Neben dem Aspekt einer schnelleren Informationsversorgung gewinnt der Gedanke der aktiven Unterstützung der Endbenutzer im Prozess der Entscheidungsfindung auf operativer und strategischer Ebene an Bedeutung. Ein aktives analyseorientiertes Informationssystem, im Sinne des Active Data Warehousing, ermöglicht eine automatisierte Benachrichtigung der Endbenutzer bei wiederkehrenden Ereignissen sowie eine automatisierte Durchführung von Entscheidungen [Sc06, 425f.]. Ziel dieses Beitrags ist die Entwicklung eines BI-Konzepts zur Erfüllung der Anforderungen von Echtzeitunternehmen und zur Umsetzung von Operational Business Intelligence. Die Entwicklung des Konzepts erfolgt dabei auf Grundlage einer Event-Driven Service-Oriented Architecture und wird sukzessiv anhand einer Ebenenstruktur vollzogen.

1 Erweiterungspotenzial von Data-Warehouse-Ansätzen zur Umsetzung von Operational Business Intelligence

Klassische DWH-Systeme als technische Umsetzung von Business Intelligence sind Lösungen, die insbesondere auf die strategische Entscheidungsfindung durch eine eingeschränkte Zielgruppe von Fach- und Führungskräften ausgerichtet sind. Auf Grundlage von überwiegend retrospektiven Analysen mittels DWH-Systemen werden strategische Unternehmensentscheidungen von den Führungskräften gefällt. Eine Unterstützung für hierarchisch niedrige, operative Ebenen im Sinne von Operational Business Intelligence ist bei den klassischen DWH-Ansätzen nicht vorgesehen [Ec07, 4f.]. Jedoch besitzt auch diese Unterstützung hinsichtlich einer Integration des Right-Time-Ansatzes hohe Relevanz, da operative Entscheidungen situationsadäquat beim Auftreten von Ereignissen getroffen werden müssen und eine Reaktion in Echtzeit ermöglicht werden sollte. Eine Verlagerung von Entscheidungen auf eine niedrigere Management-Ebene geht mit einer

Verbesserung der Qualität und der Geschwindigkeit der Entscheidungsfindung einher [CK06, 1f.; Ve07, 5]. Zudem ermöglicht die Integration einer aktiven Komponente den Endanwendern auf operativer Ebene automatisiert auf unternehmenskritische Situationen hinzuweisen oder bei einer vollständigen Integration der Geschäftsprozesse Entscheidungen autark durchzuführen (Active Data Warehousing¹) [CK06, 1f.]. In diesem Kontext wird deutlich, dass eine Anpassung der analytischen Informationssysteme auf die Bedürfnisse einer Vielzahl von Nutzern auf hierarchisch niedrigeren Ebenen notwendig ist.

Mit Hilfe eines BI-Konzepts soll den oben beschriebenen Anforderungen Rechnung getragen werden. Auf Grundlage der Architekturkonzepte der Service-Oriented Architecture (SOA) und Event-Driven Architecture (EDA) soll eine Verschmelzung der operativen und analytischen Informationssysteme erreicht werden. Dazu wird die Anordnung der essentiellen DWH-Komponenten aufgelöst und in ein generisches SOA-Modell überführt.

2 Event-Driven Service-Oriented Architecture

Eine Umsetzung der Anforderungen kann durch den Einsatz einer ereignisgesteuerten, serviceorientierten Architektur (Event-Driven Service-Oriented Architecture, EDSOA) erfolgen. Eine solche Architektur stellt Services zur Verfügung, die durch Ereignisse ausgelöst werden können. Separat betrachtet, stellen die SOA und die EDA zwei orthogonale Ansätze dar. Demnach kann eine SOA unabhängig von einer EDA implementiert werden und vice versa. Erst die Kombination beider Ansätze kann das volle Potenzial der Architekturen entfalten [Za07, 66]. Eine ereignisgesteuerte, serviceorientierte Architektur beseitigt die Schwächen der jeweiligen Architekturen und bietet den effektivsten Ansatz für eine agile IT-Infrastruktur sowie zur Lösung von Integrationsproblemen [SR05, 18].

Eine SOA ebnet einen Übergang von zuvor monolithischen und zentralisierten Systemen hin zu modularen und lose gekoppelten Informationssystemen. Eine EDA stellt eine sinnvolle Ergänzung zur Abbildung von Geschäftsereignissen und -prozessen in Echtzeit dar. Auf diese Weise können notwendige Entscheidungen in kürzerer Zeit getroffen und Verbesserungspotenziale ausgeschöpft werden. Die EDSOA bildet somit die Grundlage für ein erfolgreiches Geschäftsprozessmanagement und ist unabdingbar, um den Echtzeit-Gedanken zu realisieren [Za07, 69].

Im Kontext dieses Beitrags wird unter einer EDSOA eine SOA auf Basis eines Enterprise Service Bus (ESB) verstanden, die durch die EDA nicht-invasiv erweitert wird. Die Geschäftslogik liegt in Form von Services vor, wobei die initiiierenden und resultierenden Ereignisse der Prozesskette EDA-Ereignisse darstellen. Zur Realisierung der Ereignisverarbeitung liefert der ESB die entsprechenden Funktionalitäten [Za07, 67]. Der ESB einer EDSOA unterstützt als zentrale Kommunikationsplattform die ereignisge-

¹ Ausführliche Erläuterungen zum Active Data Warehousing finden sich in [Th01].

steuerte, asynchrone Interaktion zwischen den Services. Insbesondere verfolgt der ESB den Zweck einer garantierten und sicheren Übertragung von Nachrichten, der Bereitstellung von Adaptern zur Integration unterschiedlicher Komponenten sowie der Bereitstellung von Routing-Mechanismen zur Meldungsverteilung mittels einer umfangreichen Sammlung von Instrumenten [Li07, 136; Wh06, 1]. In einer EDSOA auf der Grundlage eines ESB stellen Anwendungen und Services abstrakte Endpunkte dar, die auf Ereignisse reagieren können.

Gemäß den vorangestellten Ausführungen können EDA und SOA als ein integratives, sich ergänzendes Konzept implementiert werden. Eine Kombination dieser Konzepte zu einer EDSOA auf Basis eines ESB steigert den Wert beider Architekturkonzepte [Wo06, 3]. Der wesentliche Nutzen einer EDSOA besteht darin, dass alle Geschäftsereignisse auf dem Event Bus vorliegen und im Rahmen von EDA-Prozessen, wie bspw. Complex Event Processing (CEP)², in Echtzeit verwendet werden können. Auf diesem Wege lässt sich eine vollständige Transparenz der Prozesse herstellen. Eine Reaktion auf eintretende Ereignisse kann unmittelbar erfolgen. Darüber hinaus wird durch die Möglichkeit, mehrere Services parallel und asynchron anzustoßen, der synchrone und sequentielle SOA-Kontrollfluss deutlicher und realitätsnaher abgebildet [Za07, 68].

3 BI-Konzept auf Grundlage einer EDSOA

An das im weiteren Verlauf vorgestellte Business-Intelligence-Konzept auf Grundlage einer EDSOA werden andere Anforderungen gestellt als an die klassischen DWH-Ansätze. Es müssen nicht nur eine größere Anzahl verschiedener Endnutzer mit analytischen Informationen versorgt, sondern ebenfalls ein höheres Datenvolumen in Real-Time/Right-Time verarbeitet und übermittelt werden. Dabei darf die Performanz der Quell- und Zielsysteme nicht eingeschränkt und die Transaktionen der operativen Informationssysteme nicht negativ beeinflusst werden. Durch den in diesem Kontext zunehmend operativen Charakter der analytischen Informationssysteme sind darüber hinaus eine hohe Ausfallsicherheit sowie geringe Wiederherstellungszeitfenster unumgänglich. Zudem gilt es, die Herausforderung zu meistern, bei bspw. komplexen multidimensionalen Abfragen Analysewerte in Real-Time/Right-Time zurückzuliefern. Letztlich sind solche Abfragen dynamisch zu formulieren, um als Services wieder verwendet werden zu können [Ec07, 19].

Die Unterteilung der traditionellen Struktur eines DWH-Systems in die Schichten Datenerfassung, Datenhaltung, Datenbereitstellung/-analyse sowie Präsentation kann im Weiteren nicht beibehalten werden. Im Kontext dieser Erweiterung können die Rollen einzelner Komponenten und somit die jeweilige Zugehörigkeit zu den Schichten nicht Bestand haben, da sich der Service- und Ereignisgedanke im Hinblick auf die Geschäftsprozessorientierung nur unzureichend in dieser Form der Architektur darstellen lässt. Infolgedessen dient als Grundlage ein generisches SOA-Architekturmodell³. In diesem

² Eine ausführliche Erläuterung des Complex Event Processing findet sich in [Lu02].

³ In Anlehnung an [Li07, 27].

Zusammenhang werden Komponenten der verschiedenen DWH-Ansätze in das SOA-Architekturmodell überführt. Anschließend werden Erläuterungen für das Zusammenspiel der einzelnen Komponenten sowie eine Darstellung möglicher Service-Typen angeführt. Abbildung 1 stellt das Business-Intelligence-Konzept auf Grundlage einer ED-SOA dar.

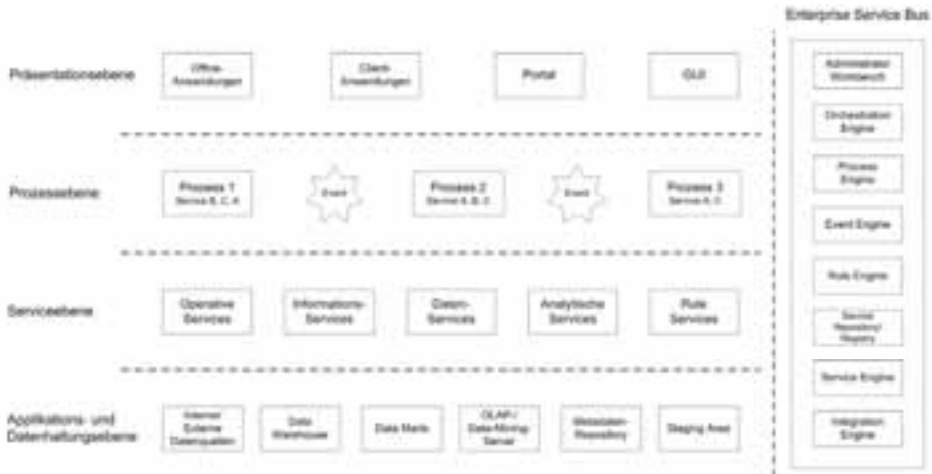


Abbildung 1: Business-Intelligence-Konzept auf Grundlage einer Event-Driven Service-Oriented Architecture

3.1 Applikations- und Datenhaltungsebene

Aufgrund der unterschiedlichen Gestaltungsziele eines Data Warehouse und einer EDSOA scheint eine Zusammenführung beider Konzepte widersprüchlich. Ein DWH dient als dauerhafter Datenspeicher, um historische Daten für Analysezwecke bereitzustellen. Zudem ist der Einsatz eines DWH auf langfristige Stabilität ausgerichtet. Mit einer EDSOA hingegen sollen Änderungen innerhalb der operativen Integration erleichtert und schneller nachvollziehbar gemacht und somit die Agilität gesteigert werden. Dieser mögliche Zielkonflikt ist insbesondere von der Beziehung zwischen operativen und analyseorientierten Informationssystemen abhängig. Die operativen Informationssysteme dienen einerseits als Datenbasis für das DWH, andererseits können die analytischen Informationssysteme aber auch als Datenquelle für die operativen Informationssysteme verstanden werden und müssen naturgemäß höheren Anforderungen gerecht werden [Sc07, 151].

Im Vergleich zu den klassischen DWH-Ansätzen ist zu erkennen, dass das DWH den Status als zentraler Kern verliert und als eine von mehreren Komponenten auf der Applikations- und Datenhaltungsebene verstanden wird. Das DWH kann somit im Rahmen der Serviceorientierung sowohl als Service-Nachfrager als auch als Service-Anbieter auftreten. Eine solche vollständige Integration ist bislang nur selten in der Praxis implementiert. Zumeist findet eine sukzessive Integration zunächst als Service-Nachfrager statt [Di07, 136f.]. Durch die Implementierung des DWH als Service-Nachfrager in eine

EDSOA werden mithilfe der zentralen ESB-Infrastruktur und den darin enthaltenen Extraktion-, Transformation-, Laden-Komponenten (ETL-Komponenten) die benötigten Daten aus den internen und externen Datenquellen in das DWH überführt. Abhängig vom Verwendungsszenario der analytischen Abfragen und der damit einhergehenden tolerierbaren Latenzzeiten kann bspw. das Konzept des Change-Data-Capture-Verfahrens (CDC-Verfahren) zur Performanzoptimierung und somit zur schnelleren Bereitstellung der Daten implementiert werden. Die benötigten Metadaten und Änderungsinformationen, um Deltas zu laden, werden mithilfe des Metadaten-Repository, das eine weitere Komponente dieser Schicht darstellt, über den ESB zur Verfügung gestellt.

Die konsequente Weiterentwicklung besteht darin, das DWH mit den aufsetzenden Analysewerkzeugen als Service-Anbieter im Sinne des Closed-Loop-Ansatzes über den ESB verfügbar zu machen und die internen und externen Datenquellen mit Analyseinformationen anzureichern. Darüber hinaus können analytische Informationen in den Ablauf von Geschäftsprozessen im Sinne von Operational Business Intelligence eingebettet werden [Di07, 137; Sc07, 152f.]. Einen ersten Schritt zur Minimierung der Latenz und somit zu einer schnelleren Bereitstellung von Analysedaten stellt die Anbindung von spezifischen Data Marts an den ESB dar, die bereits themenbezogene Daten für repetitive Abfragen vorhalten und über Services dem ESB unmittelbar zur Verfügung stehen.

Die OLAP- und Data-Mining-Server, die im klassischen DWH-System ursprünglich der Schicht der Datenbereitstellung/-analyse zugeordnet werden, stellen in diesem Konzept weitere Komponenten der Applikations- und Datenhaltungsebene dar. Ein erster Ansatz zur Zusammenführung von operativen und analytischen Daten kann über die Implementierung einer Staging Area erfolgen, um Prozesse in Echtzeit mit Analytik anzureichern. Eine temporäre Staging Area stellt eine hoch performante Datenbank zur kurzfristigen Zwischenspeicherung von operativen und analytischen Daten dar. Mithilfe von Metadaten lassen sich die eingehenden Datensätze zudem anreichern [JS07, 31].

Zusammenfassend kann konstatiert werden, dass die Applikations- und Datenhaltungsebene hierarchisch gleichwertige Anwendungen und Datenbanken umfasst, die in einer EDSOA als Backend-Systeme dienen und ihre Funktionalität den Services über den ESB zur Verfügung stellen.

3.2 Enterprise Service Bus

Der Enterprise Service Bus stellt das zentrale verbindende Element zwischen den Komponenten der verschiedenen Ebenen dar. Der ESB regelt und überwacht die Kommunikation der Services und Ereignisse. Dabei unterstützt der ESB die Datenbewirtschaftung der Systeme untereinander als asynchrone Publish-and-Subscribe-Kommunikation mithilfe der EDA-Komponente sowie als synchrone Kommunikationsform für Service-Nutzer [Di07, 138]. Aufgrund der zentralen Bedeutung zeigt Abbildung 2 den ESB im Detail.

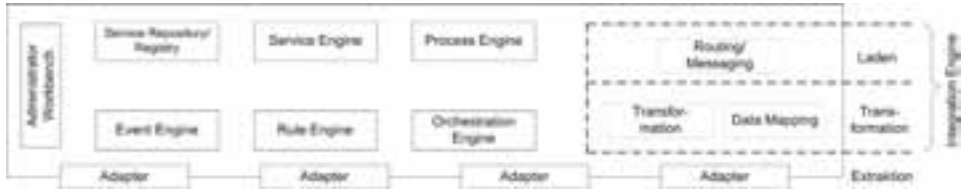


Abbildung 2: Enterprise Service Bus

Eine Interaktion der Informationssysteme und Datenbanken der Applikations- und Datenhaltungsebene untereinander ist prinzipiell multidirektional und ausschließlich über den ESB möglich.

Im Kontext von Operational Business Intelligence sind insbesondere Lösungen gefordert, die die hohen Anforderungen des Right-Time-Ansatzes erfüllen. Daher ist eine unmittelbare Datenversorgung aus den operativen Quellsystemen auf dem ESB obligatorisch, um auf Diskrepanzen im Prozessablauf sofort reagieren zu können. Damit zusätzlich analytische Abfragen für eine proaktive Geschäftsprozessüberwachung rechtzeitig verfügbar sind, müssen die operativen Daten echtzeitnah in das DWH geladen und anschließend analysiert werden. Der klassische ETL-Prozess, abhängig vom Real-Time-Anspruch, kann dahingehend als kritischer Bereich gewertet werden. Zur Realisierung müssen zunächst die ETL-Prozesse der Integration Engine im ESB schlanker formuliert werden und eine Abwendung der Verarbeitung von Massendaten zur Einzelsatzverarbeitung erfolgen, um einen kontinuierlichen Verarbeitungsstrom zu gewährleisten [Sc07a, 1]. Der beschriebene kontinuierliche ETL-Prozess der Integration Engine, als wesentliche Komponente zur Real-Time-Verarbeitung, umfasst innerhalb des ESB Adapter als Extraktionskomponente, Elemente der Transformation und des Data Mapping als Transformationskomponente sowie Routing und Messaging als Ladekomponente [JS07, 31].

Ein kontinuierlicher Verarbeitungsstrom kann mit Techniken der Datenpropagierung umgesetzt werden. In diesem Kontext ist das Verfahren der Continuous Data Integration⁴ zu favorisieren, da es zur Real-Time-Verarbeitung ereignisgesteuerte Mechanismen verwendet, die über den ESB getriggert werden können. Dadurch vollzieht sich ein Wandel von einem Pull- zu einem Push-Mechanismus. Auf diese Weise wird eine kontinuierliche und asynchrone Datenbereitstellung möglich. Die operativen Quellsysteme melden Veränderungen in den Datenbeständen als Nachricht an eine Data Change Engine, die die Änderungen sowie die geänderten Daten in eine Message Queue schreibt. Die Data Change Engine sowie die Message Queue sind in der Architektur als Routing/Messaging zusammengefasst. Mithilfe des Publish-and-Subscribe-Verfahrens der EDSOA kann die Datenintegrationsanwendung spezifische Änderungen abonnieren, diese sukzessiv verarbeiten und somit die Daten kontinuierlich in das DWH laden [LH06, 11].

Neben den zentralen ETL-Funktionen umfasst der ESB eine Orchestration, Process, Service, Rule und Event Engine. Über die Orchestration Engine werden die unterschied-

⁴ Eine ausführliche Erläuterung des Continuous-Data-Integration-Verfahrens findet sich in [TT03].

lichen Service-Abfolgen definiert, die auf der Prozessebene modelliert werden, um bspw. analytische Fragestellungen zu bearbeiten. Die eng mit der Orchestration Engine zusammenarbeitende Process Engine beinhaltet die entsprechenden Mechanismen zur Steuerung und Ausführung dieser Prozesse. Dabei ist die Process Engine insbesondere für die korrekte Abwicklung der Prozesse verantwortlich. Die Service Engine stellt Mechanismen zur Verfügung, damit Services entwickelt und integriert werden können. Die Rule Engine umfasst definierte ECA-Regeln (Event, Condition, Action) sowie analytische Entscheidungsregeln. Somit können einzelne Aktionen wie bspw. OLAP-Analysen durch definierte Szenarien automatisiert ausgeführt werden, ohne dass der Endanwender aktiv werden muss. Solche Regeln sind nicht auf Geschehnisse außerhalb des DWH-Systems anwendbar. Diesbezüglich kann die Event Engine Abhilfe schaffen. Neben den angeführten Entscheidungsregeln unterstützt der ESB mithilfe der Event Engine die Ereigniserkennung, Analyse und Verarbeitung auf dem ESB. Diese Ereignisse können unterschiedliche Services anstoßen. Die Integration von Ereignisverarbeitungsmechanismen, wie das Complex Event Processing, ermöglichen nicht nur auf Ereignisse in Echtzeit zu reagieren, sondern auch prädictiv Problemszenarien zu identifizieren und somit frühzeitig agieren zu können.

Administration- und Monitoring-Funktionalitäten zur Konfiguration aller Aktivitäten auf dem ESB sind durch die zentrale Schnittstelle Administrator Workbench implementiert. Weiterhin ist das Service Repository/Registry zur Verwaltung und Veröffentlichung der Services notwendig. Der ESB übernimmt demzufolge ebenfalls die Aufgabe eines Service-Vermittlers.

3.3 Serviceebene

Die Serviceebene umfasst unterschiedliche Service-Typen. Die Services kapseln die Funktionalitäten aus den heterogenen Systemen der Applikations- und Datenebene nach den Anforderungen der systemübergreifenden Prozesse. Die wiederverwendbaren Funktionalitäten werden über lose gekoppelte Services verfügbar gemacht. Wie zuvor beschrieben, fungiert der ESB als Service-Vermittler und veröffentlicht die definierten Services, sodass diese aufgefunden werden können. Durch die Serviceebene und die nachfolgend dargestellte Prozessebene erfolgt eine Trennung von Geschäftslogik und Prozesslogik sowie darüber hinaus von fachlichen Konzepten und der technischen Realisierung [Di07, 136]. Im Kontext klassischer DWH-Systeme geht damit ein Entwicklungsprozess von analytischen Applikationen zu analytischen Services einher. Der Serviceebene sind operative Services, Daten-Services, Informations-Services, analytische Services sowie Rule Services zugeordnet.

Operative Services können als Oberbegriff für Services verstanden werden, die eine transaktionsbezogene Geschäftslogik darstellen. Anzuführen sind bspw. das Anlegen eines neuen Kunden oder die Stornierung eines Auftrags. Daten-Services sind durch einen direkten Datenzugriff auf die Datenbestände des DWH, der Data Marts und der operativen Informationssysteme gekennzeichnet. Analytische Services können gegenüber Daten-Services umfangreiche Informationen des DWH-Systems nutzen und zu Analysen auf aggregierter Ebene verwenden. Anzuführen ist die Versorgung mit Analysen

auf dem vom Service-Nachfrager gewünschten Aggregationsniveau. Zu den analytischen Services zählen auch Analysen, die auf Real-Time-Daten angewendet und z.B. zur Prozesssteuerung eingesetzt werden [Di07, 137].

Aufgrund der zentralen Bedeutung von analytischen Services im vorgestellten BI-Konzept werden nachfolgend mögliche Ausprägungen analytischer Services dargestellt. Diese Services repräsentieren die klassischen BI-Werkzeuge, wie OLAP oder Data Mining. Eine Unterteilung von analytischen Services kann bspw. in Berichts-, Abfrage- und Analyse-Services vorgenommen werden, die jeden Daten- und Informations-Service zur Datenlieferung nutzen können und somit nicht mehr ausschließlich auf Daten aus dem DWH angewiesen sind. Weitere mögliche Service-Ausprägungen stellen Planungs- und Simulations-Services dar. Durch die Einbindung dieser Services können die Auswirkungen von einzelnen Entscheidungen simuliert und geplant werden. Somit können bspw. Individualentscheidungen geprüft und an einem übergeordneten Geschäftsziel ausgerichtet werden. Neben dieser Service-Gruppe umfassen analytische Services zudem Alerting-Services. Alerting-Services sind insbesondere im Zusammenhang mit der Ereignisorientierung unerlässlich. Bei Eintritt eines Ereignisses werden entsprechende Informationen durch das alarmauslösende Ereignis an alle Empfänger distribuiert, die diese Informationen abonniert haben, um unmittelbar Aktionen auszulösen. Darüber hinaus zählen Dashboard-Services zu den analytischen Services. Solche Services stellen Präsentationsfunktionalitäten dar, die entsprechend dem Informationsprofil des Endanwenders, z.B. als Portlet, in eine Portallösung implementiert werden. Mithilfe dieser Services werden bspw. Kennzahlen der Berichts-, Abfrage- und Analyse-Services visualisiert und personalisiert verfügbar gemacht. Letztlich umfassen die analytischen Services Key Performance Metrics (KPM) Services. Diese Services implementieren Kennzahlen, um die Leistung von Prozessen zu verwalten und zu steuern [Ma06, 16ff.].

Neben den dargelegten Daten-Services existieren sog. Informations-Services. Diese Services kombinieren neben strukturierten und unstrukturierten Daten auch operative und analytische Daten. Darüber hinaus enthält die Kategorie der Informations-Services Metadaten- und Stammdaten-Services. Die letzte Service-Kategorie in diesem Kontext bilden die Rule Services. Durch den Einsatz dieser Services werden komplexe Regeln für die Entscheidungslogik gekapselt. Auf diese Weise wird erstmals die Entscheidungslogik eine Teilmenge der Geschäftslogik. Da ein Prozess auf verschiedenen Regeln basiert und Regeln typischerweise in mehreren Prozessen verwendet werden, bietet sich eine Implementierung solcher Services im Sinne von Operational Business Intelligence an. Eine gezielte Automatisierung von menschlichen Entscheidungen kann mittels der Integration einer Entscheidungslogik via Services in Prozessen erzielt werden [Ma07, 11].

3.4 Prozessebene

Auf der Prozessebene werden die verschiedenen Geschäftsprozesse und Geschäftsregeln modelliert. Die Geschäftsfunktionalität wird unter Zuhilfenahme von Services aus den Anwendungen der Applikations- und Datenhaltungsebene nutzbar gemacht und entsprechend der Prozesslogik orchestriert. Dazu können die einzelnen Services flexibel und

dynamisch in die Prozesse eingebunden werden. Die Prozesse sind infolgedessen von den Anwendungen der Applikations- und Datenhaltungsebene unabhängig. Dementsprechend kann auf Veränderungen in den Geschäftsprozessen schnell und flexibel durch eine neue Kombination der Services reagiert werden. Durch die Orchestrierung können auch operative Geschäftsprozesse mit analytischen Informationen angereichert werden. Somit sind Geschäftsprozesse durch die Implementierung von analytischen Services plan-, überwacht- und steuerbar. Die dadurch entstandene Prozessorientierung von Business Intelligence führt zu einer Einbeziehung der an den operativen Geschäftsprozessen beteiligten Fachabteilungen und infolgedessen zu einer Ausweitung des Nutzerkreises (Operational BI) [Ma06a, 6ff.]. Darüber hinaus können bspw. Rule Services zur Automatisierung bestimmter Aktionen und unter Berücksichtigung definierter Prämissen in die Geschäftsprozesse integriert werden.

Das beschriebene Szenario berücksichtigt bisweilen keine ereignisgetriebenen Aktivitäten. Ereignisse stellen in einer EDSOA jedoch fundamentale Komponenten der Prozesskette und dementsprechend der Geschäftslogik dar. Die EDA-Komponenten können in einer SOA Szenarien erkennen sowie Ereignisse generieren und versenden. Die Abonnenten können diese entgegennehmen, auswerten und geeignete Maßnahmen parallel und asynchron in Echtzeit initiieren. Infolgedessen haben Ereignisse auslösenden und resultierenden Charakter in der Prozesskette und stoßen Services an. Services können somit nicht nur von anderen Services angestoßen werden, sondern ebenfalls von Ereignissen [Za07, 61ff.].

3.5 Präsentationsebene

Die Präsentationsebene umfasst unterschiedliche Alternativen zur Visualisierung. Durch die Verbindung von analytischen und operativen Informationssystemen in dem vorgestellten BI-Konzept sind nicht mehr nur ausschließlich die dedizierten BI-Werkzeuge für Präsentationsfunktionalitäten nutzbar. Es zeichnet sich stattdessen eine Flexibilisierung bei der Wahl der Front-End-Anwendungen ab [Ph06]. Die Darstellung der Analyseinformationen erfolgt neben Office- und Client-Anwendungen durch Portal-Lösungen. Portal-Lösungen werden über den Webbrowser verfügbar gemacht und können auf personalisierte Daten via Services zugreifen und unter Berücksichtigung entsprechender Zugriffsberechtigungen und Sicherheitsaspekte entlang der Geschäftsprozesse zur Verfügung gestellt werden. Denkbar wäre ebenfalls eine Einbeziehung von Stakeholdern durch die Distribution entsprechender Dashboard- Services zu Informationszwecken.

4 Anwendungsfall

Das vorgestellte BI-Konzept auf Basis einer EDSOA kann in der Praxis in unterschiedlichen Bereichen eingesetzt werden. Nachfolgend wird ein Anwendungsfall aus der Logistik exemplarisch aufgegriffen.

Im Bereich von Expressgut- und Schnellopaket-Zustellungen werden Lieferungen für verschiedene Kunden zu Touren zusammengefasst. In der Regel werden verspätete Zustellungen mit entsprechenden Konventionalstrafen sanktioniert. Dies erfordert eine durchgehende Überwachung der Touren mittels vordefinierter Etappenziele, um auf Probleme, wie z.B. Terminverzögerungen, unverzüglich reagieren zu können. Diese Überwachung beinhaltet eine Kontrolle, ob sich die Zustellungen im Zeitplan befinden. Im Falle einer Verzögerung im geplanten Versandablauf, die z.B. durch ein defektes Transportfahrzeug hervorgerufen wird, muss auf Grundlage einer Kosten-/Nutzenanalyse eine kurzfristige Anpassung der Tourenplanung vorgenommen werden. Hierzu sind neben operativen auch analytische Daten für eine verbesserte Entscheidungsunterstützung notwendig.

Beim Eintritt einer Terminverzögerung wird durch das auslösende Geschäftsereignis „Verzögerung bei Expresszustellung“ ein Ereignis auf dem ESB initiiert und übertragen. Ein Disponent, als Abonnent eines solchen Ereignisses, wird sofort in Kenntnis gesetzt und entsprechende Prozesse werden zur Tourenanpassung angestoßen. Die Benachrichtigung des Disponenten kann mithilfe eines Dashboard-Service erfolgen, der basierend auf einer Nachricht auf dem ESB getriggert wird. Parallel wird eine Prüfung der Daten aus einem GPS(Global Positioning System)-Tracking-System des defekten Fahrzeugs mit GPS-Daten aktuell betriebsbereiter Fahrzeuge (in regionaler Nähe) in Echtzeit ausgelöst. Über die identifizierten Fahrzeuge können operative Daten bezüglich ihrer Tour und freien Kapazitäten abgefragt werden. Weiterhin muss eine Analyse über die Art und Höhe der Konventionalstrafen und den Customer Lifetime Value (CLV) erfolgen (Kosten-/Nutzenanalyse), um eine verbesserte Grundlage für den Disponenten bezüglich der komplexen Entscheidung der Tourenanpassung/-zusammenlegung herbeizuführen. Dies kann mithilfe eines analytischen Services erreicht werden. Auf der Grundlage der gelieferten Informationen führt der Disponent eine Tourenanpassung durch. Als Ergebnis der Umplanung werden dann wiederum automatisch neue Ereignisse auf dem ESB initiiert, die z.B. einen Service anstoßen, um die von einer verspäteten Zustellung betroffenen Kunden per E-Mail zu informieren.

Die Zusammenführung der operativen und analytischen Informationen in Echtzeit, ausgelöst durch ein Ereignis, zeigt den Nutzen auf, den die Einführung eines BI-Konzepts auf Basis einer EDSOA stiften kann. Auf diese Weise können die Anforderungen von Echtzeitunternehmen zur Umsetzung von Operational Business Intelligence erfüllt werden.

5 Fazit und Ausblick

Die veränderten Anforderungen an Unternehmen ziehen Modifikationen der Systemlandschaften nach sich, die auch die analytischen Informationssysteme umfassen. Das Ziel dieses Beitrags bestand darin, durch die Entwicklung eines BI-Konzepts auf Grundlage einer EDSOA diesen veränderten Anforderungen Rechnung zu tragen. Die Implementierung einer SOA ermöglicht monolithische Systeme zu verteilten, modularen und lose gekoppelten Systemen zusammenzuführen, die eine flexible Darstellung von Geschäftsprozessen ermöglichen. Durch die Abbildung von Geschäftsereignissen im Rah-

men einer EDA besteht die Möglichkeit Geschäftsprozesse echtzeitnah auszuwerten und darzustellen. Unternehmen können sich schnell und flexibel an die veränderten Anforderungen anpassen und strategische sowie operative Entscheidungen unter Zuhilfenahme von analytischen Services treffen. Infolgedessen wird eine vollständig flexible und agile Infrastruktur realisiert. Zudem können Analysen automatisiert angestoßen und dem Endnutzer zur Entscheidungsunterstützung zur Verfügung gestellt werden. Außerdem wird eine weitreichende Automatisierung von repetitiven Entscheidungen durch eine Rule Engine ermöglicht. Es lassen sich ebenfalls mithilfe der Kombination von Complex Event Processing und einer Rule Engine prädiktive Szenarien erkennen, um darauf automatisiert reagieren zu können oder den entsprechenden Endnutzer für den manuellen Entscheidungsprozess rechtzeitig zu informieren.

Das vorgestellte Konzept stellt einen ersten Ansatz dar, den Anforderungen, wie bspw. der Echtzeitunterstützung auf operativer Ebene, die an Business Intelligence gestellt werden, gerecht zu werden. Die Überführung der Komponenten aktueller DWH-Ansätze in ein generisches EDSOA-Modell zeigt einen Paradigmenwechsel in der bislang getrennten Betrachtung von operativen und analytischen Informationssystemen auf. Vor diesem Hintergrund muss zukünftig eine spezifische Betrachtung der einzelnen Komponenten und deren Zusammenspiel sowie möglicher technischer und organisatorischer Probleme durchgeführt werden. Zukünftig werden höhere Anforderungen an Betriebskonzepte analytischer Informationssysteme gestellt, wie z.B. eine erhöhte Verfügbarkeit (24x7). Weiterhin sind Probleme beim ETL-Prozess auf Basis eines Enterprise Service Bus im Bereich der Massendatenverarbeitung zu erwarten.

Die Entwicklung eines BI-Konzepts auf Grundlage einer EDSOA zeigt, dass eine Annäherung von operativen und analytischen Informationssystemen vollzogen wird. Die Systeme treten somit zukünftig vermehrt in den Hintergrund und der Geschäftsprozess in den Mittelpunkt der Betrachtung.

6 Literaturverzeichnis

- [CK06] Chemburkar, A.; Keny, P.: Trends in Operational BI. <http://www.dmreview.com/issues/20060701/1057833-1.html>, 2006-07-01, Abruf am 2007-12-14.
- [Di07] Dittmar, C.: Latenzzeiten von Business Intelligence-Systemen. In (Gluchowski, P.; Chamoni, P.; Gersch, M.; Krebs, S.; Reinersmann, M. Hrsg.): Schlaglichter der Wirtschaftsinformatik. Verlag der Gesellschaft für Unternehmensrechnung und Controlling m.b.H., Chemnitz, 2007, S. 131-142.
- [Ec07] Eckerson, W.: Best Practices in Operational BI – Converging Analytical and Operational Processes. <http://www.tdwi.org/research/display.aspx?ID=8621>, 2007-07-01, Abruf am 2008-01-03.
- [FB06] Frank, W.; Baier, A.: Business Intelligence in SOA: Das Zusammenspiel von SOA- und BI-Konzepten – Ein Weg aus der Sackgasse. In: Javamagazin (2006) 11, S. 54-62.
- [JS07] Jonas, P.-K.; Scheuch, R.: Process Warehouse (PWH) – Metadatengetriebene Datenbewirtschaftung. In: BI-Spektrum (2007) TDWI Sonderheft, S. 29-34.

- [Li07] Liebhart, D.: SOA goes real – Service-orientierte Architekturen erfolgreich planen und einführen. Carl Hanser Verlag, München, 2007.
- [LH06] Liebhart, D.; Hänggi, F.: Does BI really needs SOA? Trivadis AG, Stuttgart, 2006.
- [Lu02] Luckham, D.: The Power of Events: An introduction to Complex Event Processing in Distributed Enterprise Systems, Addison-Wesley, Boston et al., 2002.
- [Ma06] Martin, W.: Analytics meets Enterprise SOA – Wertschöpfende Geschäftsprozesse durch Analytik. S.A.R.L. Martin, Annecy, 2006.
- [Ma06a] Martin, W.: Arcplan Information Services – Prozess- und service-orientierte Business Intelligence mit analytischen Services. S.A.R.L. Martin, Annecy, 2006.
- [Ma07] Martin, W.: SOA basierendes Geschäftsprozessmanagement. S.A.R.L. Martin, Annecy und IT Research, Sauerlach 2007.
- [Ph06] Philippi, J.: SOA trifft auf BI. <http://www.cio.de/knowledgecenter/bi/828707/>, 2006-10-25, Abruf am 2008-02-20.
- [Sc06] Schelp, J.: „Real“-Time Warehousing und EAI. In (Chamoni, P.; Gluchowski, P. Hrsg.): Analytische Informationssysteme. Springer, Heidelberg, 2006, S. 425-438.
- [Sc07] Schelp, J.: Ansatzpunkte zur Übertragung Serviceorientierter Konzepte auf das Data Warehousing. In (Gluchowski, P.; Chamoni, P.; Gersch, M.; Krebs, S.; Reinersmann, M. Hrsg.): Schlaglichter der Wirtschaftsinformatik. Verlag der Gesellschaft für Unternehmensrechnung und Controlling m.b.H., Chemnitz, 2007, S. 143-157.
- [Sc07a] Scheuch, R.: Quo Vadis Business Intelligence? Der BI-Markt erschließt Konvergenzen und entwickelt sich weiter. <http://www.apverlag.de/Online-Artikel/20070304/20070304w%20Opitz%20Consulting%20Quo%20vadis%20Business%20Intelligence.htm>, 2007-03-04, Abruf am 2007-02-19.
- [SR05] Sriraman, B.; Radhakrishnan, R.: Event Driven Architecture Augmenting Service Oriented Architectures. <http://soa.omg.org/Uploaded%20Docs/EDA/edamdasoa.pdf>, 2005-01-01, Abruf am 2008-01-28.
- [Th01] Thalhammer, T.: Active Data Warehouses – Complementing OLAP with Analysis Rules. Johannes Kepler Universität Linz, Linz, 2001.
- [TT03] Tho, N.; Tjoa, M.: Zero-Latency Data Warehousing for heterogeneous Data Sources and Continuous Data Streams. TU Wien, Wien, 2003.
- [Ve07] Ventana Research: Closing the Management Information Gap with Operational BI. <http://www.ventanaresearch.com/research/vendor.aspx?id=107>, 2007-05-01, Abruf am 2007-12-18.
- [Wh06] White, C.: A roadmap to enterprise data integration. IBM, Somers, 2006.
- [Wo06] Woolf, B.: Event-Driven Architecture and Service-Oriented Architecture. http://www.haifa.il.ibm.com/Workshops/oopsla2006/present/w06_eda_woolf.pdf, 2006-08-24, Abruf am 2008-01-28.
- [Za07] Zacharias, R.: SOA & Event Driven Architecture (EDA) – Eine perfekte Symbiose. In: Javamagazin (2007) 7, S. 60-69.

Business Intelligence Services bei T-Mobile Deutschland: Service Level Agreements und servicebezogenes Datenqualitätsmanagement zur kundengerechten Leistungserbringung

Clemens Herrmann*, Stefan-Alexander Müller**

*sd&m AG
software design & management
Berliner Straße 76
DE-63065 Offenbach
clemens.herrmann@sdm.de

**T-Mobile Deutschland GmbH
Landgrabenweg 151
DE-53227 Bonn
stefan.a.mueller@t-mobile.de

Abstract: Unter dem Begriff „IT Service Management“, welcher seit mehreren Jahren diskutiert wird und mittlerweile ein anerkanntes Gestaltungsparadigma für IT-Dienstleister ist, wird die stärkere Serviceorientierung von internen sowie externen IT-Dienstleistern verstanden. Ein wesentlicher Aspekt ist die vertragliche Vereinbarung von Leistungen in Form von Service Level Agreements (SLAs) zwischen Anbieter und Kunde mit der Zielsetzung, die Kundenorientierung zu verbessern und die Geschäftsbeziehungen zu professionalisieren. Trotz der zahlreichen Publikationen in diesem Themengebiet gehen bisher nur wenige Veröffentlichungen auf die spezifischen Herausforderungen im Kontext Business Intelligence (BI) ein. Der Artikel beschreibt daher die Einführung von SLAs für kritische BI-Leistungen bei T-Mobile Deutschland. Dabei wird insbesondere auf BI-spezifische Inhalte und Granularitäten der Leistungsvereinbarungen eingegangen. Darauf aufbauend werden die Auswirkungen auf das Datenqualitäts- und Metadatenmanagement anhand der Implementierung bei T-Mobile Deutschland aufgezeigt.

1 Motivation und Ausgangssituation

Begriffe wie IT Service Management und Serviceorientierung von IT-Organisationen werden seit mehreren Jahren sowohl in der Wissenschaft als auch in der Praxis diskutiert (vgl. [JHMO07, S. 583f.], [PBL04, S. 163]). Hierunter versteht man in erster Linie die stärkere Ausrichtung der internen und/oder externen IT-Dienstleister und deren Leistungen an den fachseitigen Kunden sowie die verstärkte Standardisierung der Zusammenarbeit mit den Fachseiten gemäß den Prozessen des IT Service Managements (vgl. [KHP04, S. 23]).

Im Zuge einer stärkeren Serviceorientierung der Business Intelligence (BI) der T-Mobile Deutschland (TMD) wird die Zielsetzung verfolgt, die Gestaltungsprinzipien der Best-Practice-Modelle des IT Service Managements im BI-Bereich schrittweise umzusetzen, um so eine hohe IT-Prozessqualität und eine kundengerechte IT-Leistungserbringung sicherzustellen. Der Fokus des vorliegenden Artikels richtet sich dabei primär auf die Einführung von Service Level Agreements für BI-Leistungen und daraus erwachsende Konsequenzen für das Datenqualitäts- und Metadatenmanagement. Andere Prozesseinführungen – wie bspw. Incident und Problem Management – sind nicht Bestandteil des Artikels, obwohl diese auch im Gesamtfokus der TMD liegen (vgl. z. B. [HWB04]).

Im Kontext Business Intelligence konnten die nachfolgenden Handlungsfelder bei der TMD identifiziert werden, die für eine stärkere Serviceorientierung sprechen:

- Die betrieblichen Leistungen der BI waren aus fachlicher Sicht nur zum Teil beschrieben. Der Zusammenhang der BI-Leistungen zu konkreten Nutzungs- bzw. Geschäftsprozessen war nur im Einzelfall erläutert. Zudem existierten im Regelfall keine Vereinbarungen über die dauerhaft nach der Entwicklung bereitzustellenden Leistungen zwischen IT und Fachseite.
- Aufgrund der unzureichenden Transparenz der BI-Leistungen war eine dedizierte Steuerung und Klassifizierung des gesamten BI-Leistungssportfolios nur eingeschränkt möglich.
- Datenqualitätsaussagen waren nur auf der Ebene Bericht oder OLAP-Würfel möglich. Aggregierte Qualitätsgesamtaussagen zur Informationsbereitstellung für einen fachseitigen Nutzungsprozess, die insbesondere auch aus Managementsicht notwendig waren, konnten nicht automatisiert erstellt werden.
- Verantwortlichkeiten für die Services der BI waren – sowohl auf Fach- als auch auf IT-Seite – teilweise nicht definiert oder fragmentiert auf unterschiedliche Projekt- und Linienorganisationsstrukturen verteilt. Dies führte teilweise zu erhöhtem und zeitraubendem Aufwand bei der Kommunikation und Koordination zur Betreuung eines Service, z.B. bei einer vorliegenden Störung und deren Behebung.

Ausgehend von diesem Handlungsbedarf wurde in einem Beratungsprojekt mit sd&m die Zielsetzung in Form einer Pilotierung verfolgt, ausgewählte kritische Services der Business Intelligence der TMD in SLAs abzubilden. Der Artikel fokussiert daher auf die Darstellung dieses Pilotprojekts, der darauf aufbauenden Anpassung des Metadatenmanagements sowie den Erfahrungen aus der Pilotphase.

Der Artikel ist wie folgt aufgebaut: In Kapitel 2 werden unterschiedliche Dimensionen des IT Service Managements erläutert und die konkrete Ausgestaltung bei TMD dargestellt. Kapitel 3 beleuchtet den Begriff IT-Service im Business Intelligence Kontext und verdeutlicht das BI-Serviceportfolio der TMD. Darauf aufbauend erfolgt in Kapitel 4 eine detaillierte Beschreibung der Inhalte von Service Level Agreements für BI-Services am Beispiel der TMD. Kapitel 5 zeigt die Erweiterungen des Metadatenmanagements auf, die notwendig sind, um relevante Informationen zu den BI-Services zu verwalten und entsprechende Qualitätsaussagen über die BI-Services zu generieren. Kapitel 6 fasst die wesentlichen Inhalte des Artikels zusammen und zeigt offene Punkte bzw. nächste Schritte auf.

2 IT Service Management

2.1 Dimensionen des IT Service Managements

Der Ausgangspunkt der durchgängigen Ausrichtung der IT-Organisation an den Anforderungen der Kunden und Nutzer auf Fachseite ist die mittlerweile vorherrschende Wahrnehmung des IT-Betriebes als Unterstützer des eigentlichen Geschäftsbetriebes. Erwartet wird eine Orientierung der IT an den Geschäftsprozessen und weniger an den eigentlichen technischen Möglichkeiten (vgl. [ITIL02, S. 9]). IT Service Management umfasst dabei zum einen die stärkere Dienstleistungsorientierung und zum anderen die damit verbundenen prozessualen Anpassungen (vgl. [BK04, S. 7f.]).

Im Rahmen von IT-Service-Management-Initiativen werden Prozesse und organisatorische Strukturen regelmäßig gemäß Best-Practice-Modellen des IT Service Managements eingeführt. Dies kann auch den Aufbau einer dedizierten organisatorischen Einheit zur Bündelung der für das IT Service Management notwendigen Kompetenzen beinhalten. In der Literatur existieren mehrere herstellerabhängige und -unabhängige Modelle, die mit unterschiedlichen Schwerpunkten und Detailtiefen das IT Service Management strukturieren (vgl. [HH03, S. 47-55], [Pr03, S. 78-93], [ZHB05, S. 13-22]). Als De-facto-Standard hat sich mittlerweile die IT Infrastructure Library (ITIL) etabliert, die inhaltlich auch den zertifizierbaren Standard ISO/IEC 20000 (vgl. [ISO05]) maßgeblich mitgeprägt hat. ITIL wurde auch im Projekt bei TMD als wesentlicher Standard herangezogen.

Zusätzlich zur Dienstleistungs- und Prozessorientierung entstand in den letzten Jahren die Dimension Architekturorientierung als neuer Entwicklungstrend (vgl. [BK04, S. 16ff.]). Dabei wird das Ziel verfolgt, durch Wiederverwendung bestehender Services oder Komponenten flexibel auf sich ändernde Kundenwünsche eingehen zu können. Unterstützt durch Gestaltungsansätze wie z. B. serviceorientierte Architekturen (SOA) sind erste Umsetzungen in der Praxis erkennbar (vgl. [LH07, S. 1644]).

Das Vorhaben bei TMD im Bereich Business Intelligence und somit dieser Artikel betrachten ausschließlich die Dienstleistungs- und Prozessdimension des IT Service Managements, wohingegen Aspekte zur IT-Architektur hier nicht im Fokus stehen.

2.2 IT Service Management bei T-Mobile Deutschland

Neben prozessualen Anpassungen zur Dienstleistungsorientierung der IT entstand im Jahr 2005 bei TMD ein eigenständiger Bereich „IT Service Management“ (ITSM) innerhalb der IT-Organisation. Hervorgegangen ist diese Organisationseinheit aus dem Bereich „Management IT Operations“, der in der Hauptsache für die Kontrolle der Vertragseinhaltung gegenüber dem konzerninternen Outsourcing-Partner T-Systems verantwortlich war. Gestaltungsparadigma des IT Service Managements ist eine Ausrichtung an ITIL sowie eine umfassende Wahrnehmung der Schnittstellenfunktion zwischen Fachabteilungen, interner Softwareentwicklung und dem Outsourcing-Partner.

Das IT Service Management lässt sich grob in Abteilungen für Service Delivery und Service Support unterteilen.

Die Bestandteile des Service Support sind, ihrem teilweise operativen Charakter folgend, an den geschäftlich orientierten IT-Bereichen eines im Massenmarkt tätigen Kommunikationsanbieters ausgerichtet: Customer Relationship Management, Rating/Billing/Payment, Enterprise Systems und Business Intelligence. Mit dieser Ausrichtung wird dem Umstand Rechnung getragen, dass trotz weitestgehender Standardisierung alle diese Bereiche unterschiedliche Anforderungen aufweisen und differenzierte Lösungen benötigen. Neben den operativen Aufgaben (z. B. Incident Management) verantwortet die Abteilung für Service Support die Kontrolle der Leistungserbringung des Outsourcingpartners im Rahmen des Service Level Managements. Die Abteilung für Service Delivery fokussiert dagegen in der Hauptsache auf Governance-Themen. Dies umfasst in erster Linie Budgetverwaltung, übergreifendes Prozessdesign und Verwaltung der Rahmenverträge mit externen Partnern.

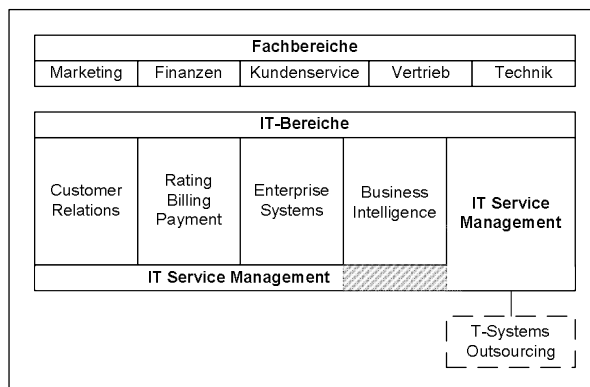


Abbildung 1: Schematischer Aufbau der Fach- und IT-Bereiche

Seitens der Organisationseinheit ITSM wurde das Projekt zur Entwicklung und Einführung von SLAs für BI durch das BI-Team der Service-Support-Abteilung verantwortet (schraffierter Bereich in Abb. 1).

3 IT-Services im Business Intelligence Kontext

3.1 Eigenschaften von IT-Services

Das im vorangegangenen Kapitel dargestellte Konzept des IT Service Managements macht die zentrale Rolle deutlich, die definierte Dienstleistungen darin einnehmen. Daher soll zunächst kurz erläutert werden, welche Eigenschaften Dienstleistungen ausmachen, um darauf aufbauend den Begriff IT-Service zu konkretisieren.

Dienstleistungen in der Betriebswirtschaftslehre unterscheiden sich von Sachleistungen durch drei konstituierende Merkmale (vgl. [Mo02, S. 89], [Be83, S. 23]):

- Die Immaterialität zielt darauf ab, dass eine Dienstleistung zumindest vor ihrer Erbringung i. d. R. nicht erfassbar und auch kaum bewertbar ist. Unter dem Gesichtspunkt der Immaterialität treten insbesondere die Notwendigkeit zur genauen Definition der erwarteten Leistung und der Ergebnismessung in den Vordergrund. Die Immaterialität ist als Merkmal nicht immer konsequent definierbar, da ein Großteil der Dienstleistungen durchaus mit materiellen Ergebnissen verknüpft ist.
- In den Prozess der Leistungserstellung werden im Regelfall externe Faktoren einbezogen. Dabei kann es sich sowohl um den Abnehmer selbst handeln als auch um das Objekt, an dem die Dienstleistung vollzogen wird.
- Die Synchronität von Leistungserstellung und -abgabe beschreibt, dass Produktion und Konsumtion einer Dienstleistung zum gleichen Zeitpunkt stattfinden („Uno-actu-Prinzip“). Im Gegensatz dazu können bei einer Sachleistung Produktion und Konsumtion zeitlich erheblich voneinander getrennt stattfinden.

Für IT-Services im Kontext IT Service Management wird eben dieser betriebswirtschaftliche Dienstleistungsbegriff zugrunde gelegt (vgl. [BHHK08, S. 62]).¹ Genauer definiert ITIL einen IT-Service als “Service provided to one or more Customers by an IT Service Provider. An IT Service is based on the use of Information Technology and supports the Customer’s Business Processes. An IT Service is made up from a combination of people, processes and technology and should be defined in a Service Level Agreement” [ITIL07, S. 301]. Dieser Begriffsauffassung wird auch in diesem Artikel gefolgt, d. h. ein IT-Service repräsentiert eine IT-Dienstleistung zur Unterstützung eines oder mehrerer Ge-

¹ Unter dem Gesichtspunkt der Kundenwahrnehmung fällt dem Merkmal der Synchronität häufig eine besondere Bedeutung zu. Die Erwartung an die Erbringung eines IT-Service orientiert sich aus Sicht des Kunden häufig an sog. Utility-Services, also Leistungen wie Strom und Wasser, die durch ein Versorgungsunternehmen erbracht werden. Die Erwartung einer ständigen Verfügbarkeit des Service kollidiert jedoch mit dem Umstand, dass sich IT-Services nicht auf Vorrat produzieren lassen. Hierdurch ergeben sich besondere Anforderungen an die Leistungsfähigkeit der IT, aber auch an die Steuerung der Kundenerwartungen. Auf dem IT-Dienstleistungsmarkt wird diese Kundenerwartung mittlerweile gezielt mit sog. Utility-Computing-Angeboten angesprochen. Schlagwörter großer Anbieter sind hier beispielsweise On-Demand Services der IBM oder Dynamic Services von T-Systems.

schäftsprozesse. Der Leistungsumfang wird vertraglich in Service Level Agreements festgelegt, deren Strukturen in Kapitel 4 eingehender erläutert werden.

3.2 Strukturierung von IT-Services bei T-Mobile Deutschland

Aus der Definition im vorangegangenen Abschnitt geht hervor, dass IT-Services in Form einer Dienstleistung direkt kundenseitige Prozesse unterstützen. Die für die Erbringung eines IT-Service notwendigen einzelnen IT-Leistungen werden oftmals auch als Services bezeichnet. Hier liegt jedoch kein betriebswirtschaftlicher, sondern ein technologieorientierter Servicebegriff zugrunde. Kernbestandteile dieser Services sind oftmals softwaretechnisch realisierte Einzelfunktionalitäten (vgl. ausführlich [BHHK08, S. 62f.]), die aus Kundensicht transparent sein sollten. Daraus ergibt sich eine hierarchische Strukturierung der Leistungserbringung, an deren Ende die IT-Services stehen. In TMD werden dabei die in Abb. 2 dargestellten Ebenen differenziert.

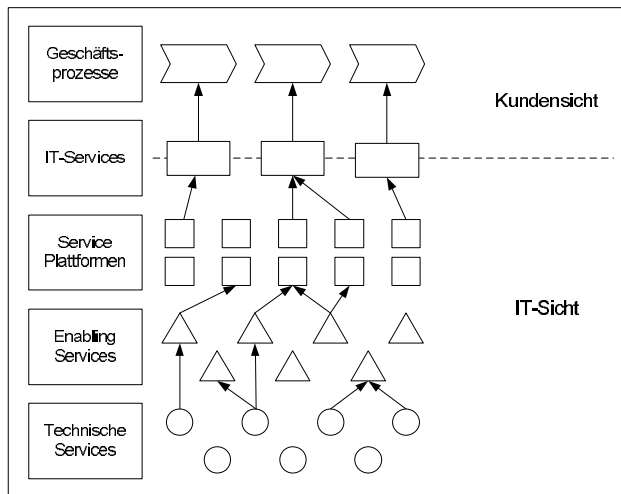


Abbildung 2: Strukturierung der Service-Ebenen bei T-Mobile Deutschland

Auf der untersten Ebene befinden sich dabei die technischen Services, d. h. Server und Infrastruktur inklusive der notwendigen Betriebsleistungen. Darauf aufbauend werden Enabling Services entwickelt, die i. d. R. jeweils eine dedizierte Aufgabe erfüllen (z. B. Vorverarbeitung von Gesprächsdatensätzen). Die Enabling Services bilden damit die Grundlage für die Bildung von Service-Plattformen, die aus mehreren unterschiedlichen Enabling Services bestehen können. Ein Beispiel dafür ist eine Plattform, die Kunden für eine Ansprache im Rahmen einer Kampagne auswählt. Hierfür sind u. a. die Kundenstammdaten und statistische Informationen über Nutzungsverhalten notwendig, die aus verschiedenen Enabling Services geliefert werden können. Auf der Ebene der IT-Services können dann eine oder mehrere Plattformen zusammengeführt werden, um einen Geschäftsprozess des Kunden zu unterstützen.

Durch diese Differenzierung von Service-Ebenen wird auch das Einkaufen bestimmter Services auf externen Märkten (Outsourcing) erleichtert. Im Regelfall weisen die Services höherer Ebenen jedoch aufgrund ihrer Kundennähe ein geringeres Standardisierungspotenzial als technische Basisleistungen auf (vgl. [BK04], S. 7). Mit sinkender Standardisierung sinkt häufig auch die Möglichkeit einer erfolgreichen Verlagerung der Leistungserbringung nach außen. Bei T-Mobile Deutschland beschränkt sich das Outsourcing daher i. d. R. auf die Ebenen der technischen und Enabling Services. Diese Art der industrialisierten Erbringung von technischen Basisleistungen wird in TMD als Factory-Approach bezeichnet. Eine Auslagerung von ganzen Unternehmensprozessen (Business Process Outsourcing) wird nicht durchgeführt, vielmehr können alternativ einzelne Geschäftsprozesse auf Konzernebene zentralisiert werden, die dann jeweils von einem sog. Shared Service Center erbracht werden.

Im Fokus dieses Artikels stehen ausschließlich die IT-Services, die zur Unterstützung der BI bezogenen Nutzungsprozesse der TMD erbracht werden. Die technischen Basisleistungen dieser IT-Services werden, wie oben beschrieben, durch den Outsourcing-partner T-Systems erbracht.

3.3 BI-Services bei T-Mobile Deutschland

In der Literatur existiert bislang keine anerkannte Definition zum Begriff BI-Service. Zwar wurden bereits einige wenige Publikationen im Kontext IT Service Management und Business Intelligence veröffentlicht (vgl. [Fi06], [SB07]), aber diese liefern keine abschließende Festlegung hinsichtlich BI-Services.

Um IT-Services im Kontext Business Intelligence näher zu erläutern, ist daher zunächst eine Beleuchtung des Begriffs Business Intelligence notwendig. Dieser bezeichnet den „IT-gestützten Zugriff auf Informationen sowie die IT-gestützte Aufbereitung und Analyse von Informationen mit dem Ziel der Unterstützung betrieblicher Entscheidungen“ [SW02, S. 439]. BI kann dementsprechend aufgefasst werden als Sammelbegriff für bestehende Analyseansätze betrieblicher Daten (vgl. [GI01, S. 5], [MS03, S. 4]).

BI-Services stellen somit eine spezifische Ausprägung von IT-Services dar, die in erster Linie relevante Informationen für Entscheider bereitstellen. Zudem stellen BI-Services die zur Nutzung der Informationen notwendigen Analysemöglichkeiten bereit. Sowohl das Informationsangebot eines BI-Service als auch die benötigten Analysefunktionen werden durch den BI-Nutzungsprozess determiniert. Dies sind i. d. R. Entscheidungsprozesse, die entweder auf strategischer, auf taktischer oder auf operativer Ebene angesiedelt sein können.

Bei TMD konnten nachfolgende Gruppen von BI-Services differenziert werden. Diese Gruppen erheben aufgrund des Pilotcharakters des Projekts keinen Anspruch auf Vollständigkeit hinsichtlich des gesamten BI-Service-Portfolios, sondern spiegeln lediglich eine Auswahl der wesentlichen BI-Services wider:

- *Unterstützung des operativen Geschäfts:* Diese Gruppe von BI-Services unterstützt direkt die operativen Kerngeschäftsprozesse des Unternehmens. Die BI-Services helfen bei der operativen Entscheidungsfindung und haben meist direkten und zeitnahen Einfluss auf die Durchführung des operativen Geschäfts. Beispiele hierfür sind die Informationsbereitstellung für das operative Kampagnenmanagement bspw. zur Kundenrückgewinnung, sowie die Informationsunterstützung zur täglichen Steuerung der Callcenter des Kundenservice bezüglich der einzusetzenden internen und externen Callcenter-Agenten.
- *Unterstützung geschäftskritischer Berichtsprozesse:* BI-Services dieser Kategorie liefern i. d. R. Informationen zur Abwicklung der Berichtsprozesse über das Kerngeschäft des Unternehmens. Oftmals sind diese Berichtsprozesse auf unternehmensexterne Empfänger ausgerichtet wie bspw. den Mutterkonzern oder die Aktionäre des Unternehmens. Die zeitgerechte und qualitativ hochwertige Bereitstellung der Informationen wird daher oftmals als kritisch angesehen. Bei TMD sind in erster Linie die BI-Services zur Unterstützung des Monats- und Jahresabschlusses dieser Gruppe zuzurechnen.
- *Standardinformationen für eine große Nutzerzahl:* Im Fokus dieser Gruppe von BI-Services steht die Bereitstellung von Standardinformationen, die im Vergleich zu den beiden vorangegangenen Kategorien als weniger kritisch eingeschätzt werden können. Im Regelfall werden die Informationen durch Analysten oder Sachbearbeiter für nicht unternehmenskritische taktische oder operative Aufgaben eingesetzt. Aufgrund der standardisierten Informationen greift oftmals ein größerer Benutzerkreis aus unterschiedlichen Fachbereichen des Unternehmens auf dieses Informationsangebot zu. Ein Beispiel eines derartigen BI-Service bei TMD sind Informationen zur Entwicklung der Neukundengewinnung und zum Verkauf von neuen Produkten.
- *Individualauswertungen mit geringer Nutzerzahl:* Diese Gruppe von BI-Services subsumiert die Informationsunterstützungen einzelner Bedarfsträger des Unternehmens, die im Regelfall nicht wiederkehrend und damit nicht standardisierbar sind. Diese BI-Services sind i. d. R. sehr heterogen, da von strategischen bis operativen Entscheidungen und von Sachbearbeiter- bis Geschäftsführungsebene sämtliche Informationsbedarfe im Rahmen dieser BI-Services bedient werden können. Beispiele solcher BI-Services bei TMD sind sog. Ad-hoc-Berichte in Form von Sonderauswertungen zu Einzelprodukten oder die Bereitstellung bestimmter „Rohinformationen“ für das Database-Marketing.

Die Auflistung der unterschiedlichen Gruppen von BI-Services verdeutlicht, dass BI-Services sehr unterschiedliche Ausprägungen annehmen können. Zur adäquaten Absicherung dieser Leistungsvielfalt zwischen der IT-Organisation und den Kunden sind geeignete vertragliche Vereinbarungen zu treffen. Dabei müssen diese Vertragswerke die verschiedenen Gruppen von BI-Service angemessen berücksichtigen. Daher wird im nächsten Kapitel, basierend auf den Erfahrungen bei TMD, eingehender erläutert, welche Inhalte dort festzuschreiben sind.

4 Service Level Agreements für BI-Services

4.1 Strukturen für Service Level Agreements

Ein SLA ist eine formell ausgehandelte, schriftlich dokumentierte Vereinbarung, die zwischen zwei voneinander unabhängigen Partnern abgeschlossen wird. Inhaltlich bezieht sich ein SLA auf die Erbringung eines IT-Service, der zur Erfüllung einer Kundenanforderung dient. Kernkomponente des SLA ist die Spezifizierung der Anforderung des Kunden in Kombination mit der Festschreibung des durch den Erbringer des Service einzuhaltenden Qualitätsniveaus (vgl. [Be05, S. 12-21], [ITIL01, S. 29f.]).²

Bei der Betrachtung von SLAs können drei Strukturformen unterschieden werden, die unterschiedliche Herangehensweisen zum Erstellen der SLAs repräsentieren (vgl. [ITIL01, S. 35]):

- *Service-based*: Ein SLA deckt einen IT-Service ab, der für alle Kunden gleich ist. Als Ausprägung kann ein Kunde diesen IT-Service beziehen oder nicht beziehen. Ein Beispiel hierfür ist die Bereitstellung von Funktionalität zum Empfang bzw. zum Versenden von E-Mails.
- *Customer-based*: Mit einer einzelnen Kundengruppe wird eine individuelle Vereinbarung getroffen, die alle von ihr genutzten IT-Services abdeckt. Grundlage ist eine individuelle Gestaltung.
- *Multi-Level*: Diese Strukturform kombiniert mehrere Ordnungskriterien zur Vereinbarung der SLAs. Auf der obersten Ebene (Corporate-Level) wird ein für alle Kundengruppen gültiger Rahmenvertrag etabliert. Ergänzend werden für einzelne Kundengruppen serviceunabhängige Regelungen festgelegt (Customer-Level). Auf der untersten Ebene erfolgt die Vereinbarung von Regelungen für individuell angepasste IT-Services je Kundengruppe (Service-Level). Ziel einer Multi-Level-Struktur ist es, den Umfang und die Komplexität der SLAs zu beschränken und Redundanzen zu vermeiden. Dadurch verringert sich auch der Aufwand für notwendige Aktualisierungen der Vereinbarungen.

Da die Kunden- und Nutzergruppe der Business Intelligence bei der TMD sämtliche Geschäftsbereiche sowie diverse Einheiten des Konzerns umfassen, bot es sich an, für die Ausgestaltung der SLAs eine Multi-Level-Struktur zu nutzen. Im Rahmen der vorliegenden Pilotierung wurden jedoch zunächst nur die Ebenen Customer- und Service-Level umgesetzt. Die Ausgestaltung der Rahmenvertragsebene (Corporate-Level) soll erst im Anschluss erfolgen, sobald eine ausreichende Zahl an kundenbezogenen SLAs vorliegt, aus denen übergreifende Regelungen ableitbar sind.

² Im ITIL-Modell finden sich ergänzend zwei weitere Vertragstypen: Operational Level Agreement (OLA) und Underpinning Contract (UC). Ein OLA ist eine nach innen gerichtete Vereinbarung auf Betriebsebene, d. h. zwischen internen Einheiten. Der UC ist ein Vertrag, der i. d. R. zwischen dem Dienstleister und einem externen Lieferanten abgeschlossen wird. Beide Vereinbarungen dienen der eigenen internen Absicherung der Leistungen, die vom Dienstleister für einen Kunden erbracht werden (vgl. [ITIL01, S. 41f.]). Im Fokus dieses Artikels steht nur das SLA.

4.2 Granularität der Leistungsbeschreibung für Business Intelligence

In der IT-Service-Management-Literatur werden diverse Beispiele für mögliche SLA-Bestandteile beschrieben. Der Inhalt eines SLA wird differenziert in vereinbarungsbezogene, dienstleistungsbezogene, managementbezogene und dokumentbezogene Elemente (vgl. ausführlich [BE05, S. 67-102]). Diese Bestandteile können als Grundgerüst für die Gestaltung einer Vorlage für SLAs genutzt werden, die dann individuell auf die jeweilige Situation zuzuschneiden ist.

Im Rahmen des Projekts bei TMD wurde eine spezifisch auf BI-Services angepasste Vorlage entwickelt. Wesentliche Elemente dabei, die insbesondere BI-Spezifika repräsentieren, sind zum einen der fachliche Bezug eines BI-Service sowie zum anderen die Kernleistung der Informationsbereitstellung eines BI-Service.

Zur Darstellung des fachlichen Bezugs ist der Nutzungsprozess zu beschreiben. Im Falle von BI-Services können dies sowohl Entscheidungsprozesse als auch direkt Geschäftsprozesse sein. Entscheidungsprozesse werden traditionell durch BI unterstützt und subsumieren alle Entscheidungsaufgaben im Unternehmen. Entscheidungsprozesse können daher auf strategischer, taktischer oder operativer Ebene angesiedelt sein. Geschäftsprozesse dagegen sind ein neueres Anwendungsgebiet für BI. Insbesondere operativ eingesetzte BI-Services unterstützen i. d. R. direkt Prozesse des Kerngeschäfts.

Ausgehend von dem Nutzungsprozess eines BI-Service, können dann die relevanten thematischen Blöcke aufgelistet werden. Hierbei sollen noch keine detaillierten Angaben zu Einzelinformationen gemacht werden. Bei TMD hat sich als hilfreich erwiesen, zu prüfen, welche analytischen Fragen im Rahmen des Prozesses zu beantworten sind. Dies erleichtert die Explikation der benötigten Informationsblöcke. Ausgehend von dem Prozess ist zudem darzustellen, welche Kritikalität der BI-Service einnimmt. Dies kann daran festgemacht werden, welche finanziellen oder qualitativen Auswirkungen ein Ausfall des BI-Service auf die Nutzer oder das Geschäft hat.

Diese grobe Beschreibung der Nutzungsprozess und des Informationsgehalts des BI-Service dient den Fachbereichen und dem Management dazu, einen Überblick zum BI-Service zu bekommen, den Nutzen des BI-Service abzuschätzen sowie beurteilen zu können, ob der BI-Service aus Kundensicht für den jeweiligen Anwendungszweck geeignet ist.

Im Anschluss an diese grobe Darstellung ist die Kernleistung des BI-Service zu definieren. Dies ist i. d. R. eine Spezifikation der durch den BI-Service bereitgestellten Informationen. Dabei sind prinzipiell unterschiedliche Granularitätsebenen denkbar, auf denen eine Festlegung erfolgen kann (vgl. Abb. 3):

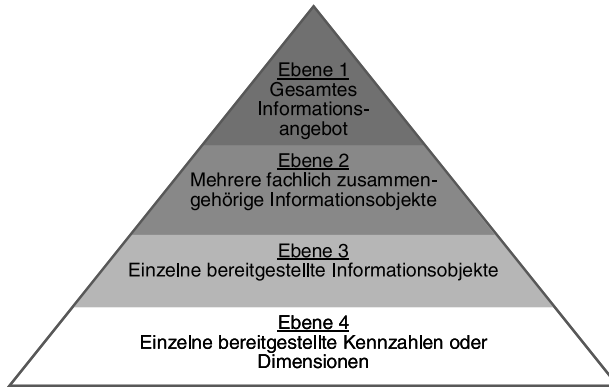


Abbildung 3: Potenzielle Granularitätsebenen zur Festschreibung von BI-Leistungen in SLAs

- *Ebene 1* repräsentiert das gesamte Informationsangebot, welches durch die Business Intelligence bereitgestellt wird.
- *Ebene 2* beschreibt fachlich zusammengehörige Informationsobjekte innerhalb des gesamten Informationsangebots. Unter einem Informationsobjekt wird dabei grundsätzlich eine festgelegte Menge von Informationen verstanden, die als eine Einheit angeboten werden kann. Dies kann je nach Ausprägung ein Bericht mit festgelegten Kennzahlen, ein OLAP-Würfel mit einer festgelegten Kombination aus Dimensions- und Faktendaten oder eine Tabelle mit definierten Attributen sein.
- *Ebene 3* beinhaltet einzelne dieser oben beschriebenen Informationsobjekte.
- *Ebene 4* spiegelt einzelne Dimensions- oder Faktendaten wider, aus denen die Informationsobjekte aufgebaut sind.

Teilweise implizieren bestimmte Typen von Informationsobjekten noch einen definierten Umfang an Analysefunktionalität, wie z. B. OLAP-Würfel, bei deren Zugriff im Regelfall Operationen (z. B. Slicing, Dicing, Drill-up, Drill-down) auf den angebotenen Informationen ausgeführt werden können. Diese Analysefunktionalität stand im Pilotierungsprojekt nicht im Fokus und wurde nicht als Leistungsbestandteil in den SLAs aufgenommen. Gegebenenfalls ist dies bei der Betrachtung weiterer BI-Services bzw. BI-Servicegruppen als Vertragsgegenstand bei der Leistungserbringung zu ergänzen.

Aus Sicht der vertraglichen Leistungsvereinbarung ist grundsätzlich zu klären, auf welcher dieser vier Granularitätsebenen die Leistungsbeschreibung im SLA vorzunehmen ist. Bei TMD wurde die Ebene 2 als adäquate Granularitätsebene festgelegt, d. h. in den SLAs für BI-Services werden die durch den BI-Service für einen Nutzungsprozess bereitgestellten Informationsobjekte festgeschrieben. Einzelne Kennzahlen oder Dimensionen der Informationsobjekte werden nicht in den SLAs erfasst. Dies dient zum einen der Reduktion von Pflegeaufwand, so dass kleinere Anpassungen der Informationsobjekte keine Aktualisierungen der SLAs notwendig machen. Zum anderen werden Redundanzen zu den im Rahmen der Entwicklung erstellten Dokumenten vermieden, die insb. die softwaretechnische Realisierung im Detail beinhalten, jedoch aus betrieblicher Sicht weniger von Interesse sind.

Andere Granularitätsebenen zur Leistungsbeschreibung können gewählt werden, wenn BI-Services beispielsweise extern erbracht werden. In diesem Fall liegt i. d. R. ein höherer Sicherheitsbedarf beider Vertragsparteien hinsichtlich des zu erwartenden Leistungsumfangs vor, so dass feingranularere Leistungsbeschreibungen sinnvoll erscheinen. Im Falle von TMD werden die BI-Services jedoch ausschließlich konzernintern genutzt. Dadurch ist eine ausreichend große Vertrauensbasis zwischen Leistungsersteller und Leistungsnehmer vorhanden, so dass detailliertere Leistungsbeschreibungen als Ebene 2 nicht notwendig sind.

Nach der Erläuterung der inhaltlichen Dimension von SLAs für BI-Services soll darauf aufbauend im nächsten Abschnitt auf die festzuschreibenden Qualitätsparameter eingegangen werden.

4.3 Qualitätsparameter für die Leistungsbeschreibung von Informationsobjekten

Die Qualitätsparameter einer Leistungsbeschreibung legen fest, welches Qualitätsniveau die einzelnen Leistungsbestandteile erfüllen müssen. Die zentralen Leistungsbestandteile von BI-Services stellen die Informationsobjekte dar. Für diese Berichte, OLAP-Würfel oder Tabellen kann die Qualität hinsichtlich verschiedener Merkmale festgelegt werden. Bei TMD haben sich im Laufe des Pilotprojekts die vier Merkmale Aktualität, Vollständigkeit, Korrektheit und Leistungsverhalten als die aus Kundensicht wichtigsten und praxistauglichsten Qualitätseigenschaften herauskristallisiert:³

- Die Aktualität eines Informationsobjekts beschreibt, welchen Aktualitätsstand die zuletzt dem Objekt hinzugefügten Informationen haben. Beispielsweise besitzen viele OLAP-Würfel bei TMD die Aktualität „Informationen des Vortages“, d. h. gemessen an dem Zeitpunkt der „Entstehung“ der Daten im Quellsystem sind diese im entsprechenden Informationsobjekt mit einer Verzögerung von einem Tag sichtbar. Zur Aktualität zählt auch der Bereitstellungszeitpunkt eines Informationsobjekts. Hierunter wird das genaue Datum und die Uhrzeit verstanden, an dem das Informationsobjekt dem Anwender mit den entsprechend aktuellen Daten zur Nutzung zur Verfügung steht. Der Bereitstellungszeitpunkt wird in hohem Maße auch durch den Nutzungsprozess determiniert.
- Das Merkmal Vollständigkeit eines Informationsobjekts bezeichnet den erwarteten fachlichen Umfang der Informationen in einem Objekt. Dies bezieht sich in erster Linie auf die erwartete Menge an Datensätzen. Die Vollständigkeit kann sich jedoch auch auf die Attribute von Datensätzen beziehen. Bei der Erstellung der SLAs bei TMD stellte sich heraus, dass oftmals keine „echte“ Vollständigkeit im Vergleich zu real stattgefundenen Ereignissen gefordert wurde, sondern die Vollständigkeit gegenüber den Daten im Quellsystem ausreichend war.

³ In der Literatur existieren zahlreiche Kriterienkataloge und Ordnungsrahmen zur Untergliederung der Datenqualität (vgl. ausführlich [He02, S. 69-95]).

- Die Korrektheit von Informationsobjekten spiegelt die fachliche Richtigkeit von Informationen wider. Dabei können im Regelfall nur Plausibilitätsaussagen getroffen werden, da eine eindeutige Korrektheit sich oftmals der Messbarkeit entzieht. Bei TMD stellte sich die Explikation von Korrektheitskriterien als aufwendig heraus und es konnten nur in einzelnen Fällen initial Indikatoren für die Korrektheit identifiziert werden. Dabei ist festzulegen, gegen welche Informationen eine Messung der Korrektheit erfolgen soll. Dies kann zum einen über definierte Schwellenwerte erfolgen, anhand derer die Plausibilität der Daten geprüft wird. Zum anderen können auch die Informationen mehrere Informationsobjekte abgeglichen werden und aufgrund der „Stimmigkeit“ die Korrektheit plausibilisiert werden.
- Das Leistungsverhalten im Sinne der Geschwindigkeit der Rückmeldung an den Nutzer ist abhängig von der Art des Informationsobjekts und spielt nur bei interaktiven Analysemöglichkeiten, wie z. B. OLAP Würfeln oder parametrisierbaren Berichten, eine Rolle. Hierbei können als Qualitätsmerkmale die Zeiten zur Erstellung, zum Zugriff oder zum Manipulieren der Informationen herangezogen werden.

Diese Qualitätsmerkmale sind je Informationsobjekt mit den Vertragsparteien abzustimmen und in das SLA aufzunehmen. Für jedes Merkmal sind dann entsprechende Erreichungsgrade zu vereinbaren, mit denen festgelegt wird, wie häufig bzw. wie gut die Qualitätsparameter durch den Leistungsersteller einzuhalten sind. Da bestimmte Informationsobjekte auch in mehreren BI-Services genutzt werden können, ist es denkbar, dass für ein Informationsobjekt unterschiedliche Qualitätsfestlegungen in verschiedenen SLAs getroffen werden. Auch ist auf die Messbarkeit und Nachvollziehbarkeit der Festlegungen zu achten, damit die Erreichung der vereinbarten Leistungen geprüft werden kann. Im nächsten Kapitel wird daher aufgezeigt, wie bei TMD regelmäßig zu den BI-Services auf Basis der Metadaten berichtet wird.

5 Notwendige Metadaten für BI-Services

5.1 Metadaten-Repository

Zur Überwachung der im vorangegangenen Kapitel erläuterten Festlegungen für die Informationsobjekte je BI-Service und zur Erstellung servicebezogener Statusberichte ist eine geeignete Werkzeugunterstützung notwendig. Dadurch sind zum einen die vertraglich dokumentierten Vereinbarungen strukturiert verwaltbar und auswertbar. Zum anderen kann die Erreichung der vereinbarten Service Levels automatisiert geprüft und kommuniziert werden.

Derartige Qualitätsinformationen zu Informationsobjekten werden im Regelfall in Form von Metadaten gehalten (vgl. [HH02, S. 101]). Bei der TMD wurde daher die Entscheidung getroffen, das bestehende zusammen mit sd&m implementierte Metadaten-Repository zu erweitern, da dort bereits einige für BI-Services relevante Metadatenentitäten vorhanden waren. Abb. 4 zeigt einen Ausschnitt des Metadatenmodells mit den durchgeführten servicebezogenen Erweiterungen und den Beziehungen zu den bestehenden Entitäten, die nachfolgend erläutert werden.

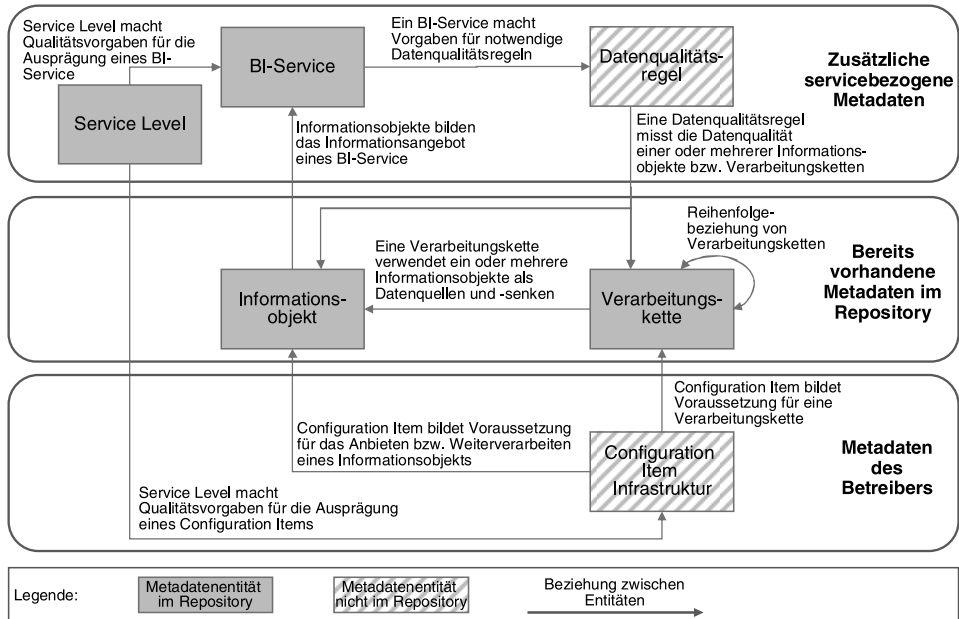


Abbildung 4: Ausschnitt des Datenmodells der Metadaten für BI-Services bei TMD

Folgende Metadaten-Entitäten waren bereits vor der Erweiterung im Repository vorhanden:

- *Entität Informationsobjekt:* Informationsobjekte bilden den Anfangs- bzw. Endpunkt von Verarbeitungsketten. Informationsobjekte stehen stellvertretend für Daten oder Informationen, die verarbeitet und dauerhaft gespeichert werden.
- *Entität Verarbeitungskette:* Eine Verarbeitungskette bildet eine abgeschlossene Einheit von Verarbeitungsschritten zur Transformation eines persistenten Datenzustands in einen neuen. Dabei dienen ein oder mehrere Informationsobjekte als Quelle und Ziel der Datenverarbeitung. Verarbeitungsketten stehen in einer Reihenfolgebeziehung zueinander bzw. werden bei bestimmten Startbedingungen ausgeführt.

Zur Abbildung und Verwaltung von BI-Services wurde das Metadaten-Repository um die folgenden zwei Entitäten erweitert:

- *Entität BI-Service:* Ein BI-Service hat einen direkten Bezug zum Anwender und beschreibt ein in sich geschlossenes Informationsangebot für einen bestimmten Geschäftsprozess. Die sich aus dem Geschäftsprozess ergebenden Anforderungen an das Informationsangebot legen den zu wählenden Service Level fest, welcher das Qualitätsniveau des BI-Service beschreibt.
- *Entität Service Level:* Ein Service Level ist eine Zusammenfassung einzelner Qualitätsparameter zu einem bestimmten Qualitätsniveau, das einem Service zugeordnet werden kann. Ein konkreter Service Level ergibt sich im Regelfall aus den Anforderungen an einen spezifischen Service. Die zu spezifizierenden Qualitätsparameter können in Abhängigkeit des Service variieren. Einheitliche Service Levels für die BI-Services konnten aufgrund des Pilotcharakters des Projekts bei TMD bisher noch nicht definiert werden. Dies ist jedoch geplant und wurde daher bereits im Datenmodell vorgesehen. Daher ist diese Entität zurzeit in Form einer textuelle Beschreibung der Service Levels implementiert.

Zusätzlich sind zur Messung der geforderten Datenqualität noch entsprechende Regeln implementiert, die zurzeit im Rahmen einer Pilotumsetzung erprobt werden. Daher sind diese noch nicht in das Metadaten-Repository integriert, sollen aber mittelfristig ebenfalls aufgenommen werden:

- *Entität Datenqualitätsregel:* Eine Datenqualitätsregel umfasst die zur Prüfung eines bestimmten Datenqualitätsaspekts eines Informationsobjekts oder einer Verarbeitungskette notwendigen Angaben. Eine Datenqualitätsregel operationalisiert damit die für einen BI-Service festgelegten Qualitätsparameter für Informationsobjekte.

Des Weiteren verwaltet der Betreiber T-Systems Informationen über die Infrastruktur in Form von Configuration Items. Diese sind ebenfalls bislang noch nicht in das Metadaten-Repository integriert, sondern werden durch den Betreiber selbst verwaltet und gepflegt. Die Metadaten der Infrastruktur sind notwendig, um z. B. gesicherte Aussagen über die mögliche Verfügbarkeit oder Performance von Informationsobjekten machen zu können:

- *Entität Configuration Item Infrastruktur:* Ein Configuration Item beschreibt ein Element der Infrastruktur, wie z. B. unterschiedliche Datenbank- oder Applikationsserver sowie betriebliche Leistungsvereinbarungen (vgl. [ITIL00, S. 122ff.]). Aktuell werden diese Metadaten in unterschiedlichen Granularitätsstufen für TMD durch den Betreiber T-Systems verwaltet und gepflegt. Aus Sicht der TMD ist insbesondere die oberste Granularitätsebene relevant. Für Configuration Items selbst gelten analog zu BI-Services entsprechende Service Levels, die deren Leistungsparameter festlegen. Zudem können Configuration Items den Informationsobjekten bzw. Verarbeitungsketten zugeordnet werden, deren „Funktionieren“ durch das Configuration Item ermöglicht wird.

Für die obigen Entitäten werden sowohl Stammdaten (wie z. B. Bezeichnungen und Verantwortliche) als auch Bewegungsdaten (wie z. B. Aktualitäten von Datenbeständen und Datenqualitätsinformationen) im Metadaten-Repository gehalten. Eine ausführliche Erläuterung dieser Attribute soll aus Platzgründen an dieser Stelle nicht erfolgen. Vielmehr soll im nächsten Abschnitt dargestellt werden, welche Berichte zu BI-Services auf Basis dieser Metadaten bei TMD angeboten werden.

5.2 Metadatenberichte

Für BI-Services gibt es im Wesentlichen zwei zusätzliche Metadatenberichte. Der Überblicksbericht liefert eine Auflistung sämtlicher vorhandener BI-Services mit einer Kurzbeschreibung und dem aktuellen, servicebezogenen Status der Informationsobjekte des jeweiligen BI-Service in Form einer Ampel. Dieser Bericht dient dem Management in erster Linie dazu, möglichst rasch den aktuellen Zustand der Informationsbereitstellung der BI beurteilen zu können.

Eine detailliertere Sicht auf BI-Services bietet der zweite Bericht, der je BI-Service ausführlich sowohl Stamm- als auch aktuelle Zustandsdaten beinhaltet. Abb. 5 zeigt einen fiktiven Bericht zu einem BI-Service „Informationsbereitstellung Monatsabschluss“. Die Berichtsdarstellung basiert auf den fachkonzeptionellen Erweiterungen für das Metadaten-Repository, da die Umsetzung der Berichte zurzeit noch im Gange ist. Dieser Bericht fokussiert in erster Linie auf die Zielgruppe der Serviceverantwortlichen, die Informationen zum Detailstatus eines BI-Service benötigen oder denen die Pflege der Metadaten zu einem BI-Service obliegt.

1

BI-Service

Name

Beschreibung

Fachlich verantwortlich

Technisch verantwortlich (IT-BI)

Technisch verantwortlich (IT-SM)

Gesamtampelstatus

Anzahl Informationsobjekte grün

Anzahl Informationsobjekte gelb

Anzahl Informationsobjekte rot

Audit-Zeitpunkt

Informationsbereitstellung Monatsabschluss

Bereitstellung der Informationen für sämtliche Phasen des Monatsabschlusses der TMD

verantwortlicher1@t-mobile.de

verantwortlicher2@t-mobile.de

verantwortlicher3@t-mobile.de

gelb

3

1

0

16.01.2008 11:32

Aktuelles Datum

16.01.2008 17:16 Uhr

2

Service-Level

Service-Level Name	Service-Level Beschreibung
Support TSES	Anwesenheits-, Reaktions- und Bearbeitungszeiten für den Support durch T-Systems.
Support TMD	Anwesenheits-, Reaktions- und Bearbeitungszeiten für den Support durch T-Mobile.
Verfügbarkeiten und Auszeiten	Generelle Verfügbarkeiten und Ausfallzeiten je Applikation, die mit T-Systems vereinbart sind.

3

Enthaltene Informationsobjekte

Name	Beschreibung	Schlechteste Datenqualität im erwarteten Zeitraum	Ampelstatus
Cube: Budget	Budgetauswertungen	Vollständig	grün
Cube: Bestand	Teilnehmerbestand	Vollständig	grün
Cube: MORPU	Monthly Revenue per User (MORPU)	Partiell	gelb
Cube: Vertrag	Vertragsverlängerungen	Vollständig	grün

4

Dokumente zum BI-Service

Name	Beschreibung	Version	Versionsdatum
BI-Service Monatsabschluss	BI-Servicevereinbarung zum Monatsabschluss	8.1	03.01.2008

5

Aktualität und Herleitung des Status der enthaltenen Informationsobjekte

Name	Erwartete Aktualität (Aktualisierungszeitplan)	Tatsächliche Aktualität (Verspätung)	Minimale Verspätung für Ampelstatus "gelb"	Minimale Verspätung für Ampelstatus "rot"
Cube: Budget	T-1 (täglich um 8:00 Uhr)	T-1		0 Tage 0 Stunden
Cube: Bestand	T-1 (täglich um 8:00 Uhr)	T-1		0 Tage 0 Stunden
Cube: MORPU	T-1 (täglich um 11:00 Uhr)	T-2 (6 Stunden 17 Minuten)	0 Tage 0 Stunden	
Cube: Vertrag	W-1 (Mittwochs um 16:00 Uhr)	W-1	0 Tage 0 Stunden	

Abbildung 5: Fiktiver Bericht zu einem BI-Service basierend auf dem Metadaten-Repository (vgl. [DR08, S. 139])

Der Bericht setzt sich aus mehreren Teilen zusammen:

1.
- Der Berichtskopf umfasst die Überblicksinformationen zum BI-Service, wie z. B. die Service-Beschreibung und die fachlichen sowie technischen Verantwortlichen. Zudem werden die aggregierten Statusinformationen der Informationsobjekte des BI-Service dargestellt.
2.
- Der nächste Berichtsteil beinhaltet die für den BI-Service gültigen Service Levels. Im Regelfall wird hier auf weiterführende Dokumente wie z. B. Supportvereinbarungen oder betriebliche Verträge verwiesen.

3. In diesem Teil des Berichts wird der fachliche Kern des BI-Service dargestellt. Hier werden die dem BI-Service zugehörigen Informationsobjekte mit ihrem jeweiligen Datenqualitäts- und Aktualitätsstatus in Bezug auf den BI-Service aufgelistet. Die Datenqualität kann in den Ausprägungen fehlerhaft, fehlend, partiell und vollständig beschrieben werden, so dass Aussagen zur Vollständigkeit und Korrektheit der Informationsobjekte möglich sind. Das Leistungsverhalten ist in dem Bericht nicht enthalten, da die hierfür notwendigen Metadaten durch den Betreiber T-Systems verwaltet werden und aktuell nicht Bestandteil des Metadaten-Repositories sind.
4. Der nächste Teil des Berichts dient der Ablage der vertraglichen Dokumente zum BI-Service. Im Regelfall wird hier das SLA verwaltet.
5. Zuletzt wird im Bericht angegeben, welche Aktualitäten für die Informationsobjekte aus Sicht des BI-Service festgelegt sind („Erwartete Aktualität“) und wie ggf. eine aktuell vorliegende Verspätung bei der Bereitstellung eines Informationsobjekts ausfällt („Tatsächliche Aktualität“). Zudem werden die Schwellenwerte zur Bestimmung des Ampelstatus gelb und rot je Informationsobjekt angezeigt, die in Form von Verspätungen gegenüber der erwarteten Aktualität verwaltet werden. Der Gesamtstatus eines BI-Service, der im Kopfbereich des Berichts dargestellt wird, entspricht dem schlechtesten Ampelstatus der Informationsobjekte des BI-Service. Aktuell bezieht sich dieser Ampelstatus ausschließlich auf die Aktualität, da diese bei TMD als das wichtigste Qualitätsmerkmal betrachtet wird.

6 Zusammenfassung und Ausblick

Der Artikel konkretisiert den Begriff IT-Service im Kontext Business Intelligence. Sogenannte BI-Services stellen eine spezifische Ausprägung von IT-Services dar, die in erster Linie Entscheider mit Informationen und ggf. Analysefunktionalität versorgen und damit einen bestimmten Entscheidungs- oder Geschäftsprozess unterstützen. Die im Rahmen des Pilotprojekts bei TMD vereinbarten SLAs charakterisieren BI-Services anhand des Einsatzzwecks, der unterstützten Nutzungsprozesse, der inhaltlich notwendigen Themenbereiche sowie der Servicekritikalität. Darüber hinaus werden die bereitzustellenden Informationsobjekte definiert und anhand der Qualitätsparameter Aktualität, Vollständigkeit, Korrektheit und Leistungsverhalten beschrieben. Zur Pflege und zur Überwachung der SLAs werden das um servicebezogene Entitäten erweiterte Metadaten-Repository und darauf aufbauende Statusberichte je BI-Service genutzt. Aufgrund des Pilotcharakters des Projekts können noch keine gesicherten Aussagen über das gesamte BI-Serviceportfolio der TMD und einen generellen BI-Service-Katalog getroffen werden.

Die bei TMD gemachten Erfahrungen sind auch in anderen Unternehmenskontexten anwendbar. Die Herleitung und Festlegung des Begriffs BI-Service ist als Grundlage und Ausgangsbasis für unternehmensspezifische Ausprägungen nutzbar. Die dazu korrespondierenden SLAs sind ebenfalls auf andere BI-Kontexte übertragbar, da diese sich in erster Linie auf ein typisches BI-Leistungsangebot und weniger auf unternehmensspezifische Aspekte bei TMD beziehen. Dies trifft insbesondere auch auf die verschiedenen Granularitätsebenen der Informationsbereitstellung (vgl. Abb. 3) als Grundlage der Leistungsfestschreibung zu.

7 Literaturverzeichnis

- [Be83] Berekhoven, L.: Der Dienstleistungsmarkt in der Bundesrepublik Deutschland: theoretische Fundierung und empirische Analyse, Band 1, Göttingen, 1983.
- [Be05] Berger, T.G.: Konzeption und Management von Service-Level-Agreements für IT-Dienstleistungen, Dissertation der Technischen Universität Darmstadt, Darmstadt, 2005.
- [BK04] Böhmann, T.; Krcmar, H.: Grundlagen und Entwicklungstrends im IT-Servicemanagement. In: HMD – Praxis der Wirtschaftsinformatik, o. Jg., Nr. 237, 2004; S. 7-21.
- [BHHK08] Buhl, H.U.; Heinrich, B.; Henneberger, M.; Krammer, A.: Service Science. In: Wirtschaftsinformatik, 50. Jg., Nr. 1, 2008; S. 60-65.
- [DR08] Dehnen, W.; Rickert, M.: Detailed System Specification AURORA Metadaten. Bonn 2008. (intern – nicht veröffentlicht)
- [Fi06] Finger, R.: BI-Betriebsmodelle auf dem Prüfstand. In: BI-Spektrum, o. Jg., Nr. 2, 2006; S. 21-25.
- [Gl01] Gluchowski, P.: Business Intelligence: Konzepte, Technologien und Einsatzbereiche. In: HMD – Praxis der Wirtschaftsinformatik, o. Jg., Nr. 222, 2001; S. 5-15.
- [He02] Helfert, M.: Proaktives Datenqualitätsmanagement in Data-Warehouse-Systemen, Logos, Berlin, 2002.
- [HH02] Helfert, M.; Herrmann, C.: Proactive Data Quality Management for Data Warehouse Systems – A Metadata based Data Quality System. In: Proceedings of 4th International Workshop on Design and Management of Data Warehouses (DMDW 2002) in conjunction with CAiSE 2002, Toronto 2002, University of Toronto Bookstores, Toronto, 2002; S. 97-106.
- [HH03] Hochstein, A.; Hunziker, A.: Serviceorientierte Referenzmodelle des IT-Managements. In: HMD – Praxis der Wirtschaftsinformatik, o. Jg., Nr. 232, 2003; S. 46-56.
- [HWB04] Hochstein, A.; Wetzel, Y.; Brenner, W.: Fallstudie: ITIL-konformer Service Desk bei T-Mobile Deutschland. In: HMD – Praxis der Wirtschaftsinformatik, o. Jg., Nr. 237, 2004; S. 32-42.
- [JHMO07] Johnson, M.W.; Hately, A.; Miller, B.A.; Orr, R.: Evolving standards for IT service management. In: IBM Systems Journal, 46. Jg. Nr. 3, 2007; S. 583-597.

- [KHP04] Kemper, H.-G.; Hadjicharalambous, E.; Paschke, J.: IT-Servicemanagement in deutschen Unternehmen – Ergebnisse einer empirischen Studie zu ITIL. In: HMD – Praxis der Wirtschaftsinformatik, o. Jg., Nr. 237, 2004; S. 22-31.
- [LH07] Legner, C.; Heutschi, R.: SOA Adoption in Practice – Findings from Early SOA Implementations. In: Proceedings of the 15th European Conference on Information Systems (ECIS), St. Gallen, 2007; S. 1643-1654.
- [MS03] Messerschmidt, H.; Schweinsberg, K.: OLAP mit dem SQL-Server: Eine Einführung in Theorie und Praxis, dpunkt, Heidelberg, 2003.
- [Mo02] Morschett, D.: Retail Branding und integriertes Handelsmarketing, DUV, 2002.
- [ITIL00] o. V.: Service Support (IT Infrastructure Library), The Stationery Office, London, 2000.
- [ITIL01] o. V.: Service Delivery (IT Infrastructure Library), The Stationery Office, London, 2001.
- [ITIL02] o. V.: Application Management (IT Infrastructure Library), The Stationery Office, London, 2002.
- [ITIL07] o. V.: Service Design (IT Infrastructure Library), The Stationery Office, London, 2007.
- [ISO05] o. V.: ISO/IEC 20000-1:2005, Information Technology – Service Management – Part 1: Specification, International Organization for Standardization, Genf, 2005.
- [PBL04] Potgieter, B.C.; Botha, J. H.; Lew, C.: Evidence that use of the ITIL framework is effective. In: Proceedings of the 17th Conference of the National Advisory Committee on Computing Qualifications (NACCQ), 2004; S. 160-167.
- [Pr03] Probst, C.: Referenzmodell für IT-Service-Informationssysteme, Logos, Berlin, 2003.
- [SB07] Schultze, A.; Brandes, H.: Service Level Agreements für BI. In: BI-Spektrum, o. Jg., Nr. 4, 2007; S. 24-27.
- [SW02] Strauch, B.; Winter, R.: Stichwort: Business Intelligence. In: (Bellmann, M.; Krcmar, H.; Sommerlatte, T., Hrsg.): Praxishandbuch Wissensmanagement – Strategien, Methoden, Fallbeispiele, Symposium, Düsseldorf, 2002; S. 439-448.
- [ZHB05] Zarnekow, R.; Hochstein, A.; Brenner, W.: Serviceorientiertes IT-Management: ITIL-Best-Practices und -Fallstudien, Springer, Berlin et al., 2005.

Governance in der Informationslogistik am Beispiel eines Energieversorgers

Barbara Dinter*, Gerrit Lahrmann*, Dirk Meyer**, Moritz Schmalz*

*Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedberg-Strasse 8
CH-9000 St. Gallen

[barbara.dinter | gerrit.lahrmann | moritz.schmalz]@unisg.ch

**E.ON AG
E.ON-Platz 1
D-40479 Düsseldorf
dirk.meyer@eon.com

Abstract: Die Informationslogistik ist ein wichtiger Bestandteil betrieblicher Informationssysteme. Um allerdings ihr volles Potenzial entfalten zu können, muss die Informationslogistik (IL) eng an den Erfordernissen der Fachbereiche ausgerichtet werden. Governance in der Informationslogistik kann dazu beitragen, dieses Alignment herzustellen. Der Beitrag entwickelt, ausgehend von der Definition der Informationslogistik und von einer gängigen Auffassung von IT-Governance, ein Verständnis von IL-Governance. Die Ausgestaltung der IL-Governance wird in aufbau- und ablauforganisatorischer Hinsicht beleuchtet und die Gestaltungsobjekte der IL-Governance werden detailliert. Eine Fallstudie zeigt am Beispiel des Energieversorgers E.ON die Umsetzung der IL-Governance in der unternehmerischen Praxis.

1 Einführung

Analytische Informationssysteme sind mittlerweile zu einem unverzichtbaren Bestandteil der Entscheidungsfindungsprozesse in Unternehmen geworden [Wi00a]. Nicht mehr der initiale Aufbau analytischer Informationssysteme steht im Fokus des Interesses, sondern Fragestellungen des Betriebs und der kontinuierlichen Weiterentwicklung analytischer Systeme. Sie erfordern eine umfassende Gesamtsicht in Bezug auf alle Initiativen und Projekte in diesem Umfeld sowie einen langfristigen Planungs- und Investitionshorizont – insbesondere, da Entscheidungsprozesse, Investitionen in IT und Informationslogistik sich signifikant auf den Unternehmenserfolg auswirken [De05; DS96; KD04]. Das St. Galler Konzept der Informationslogistik greift eben diese Potenziale auf: Es erweitert bestehende Konzepte um den Charakter einer bereichsübergreifenden, an fachlichen Zielen orientierten Informationsversorgung [DW08a].

Verschiedene Untersuchungen messen Konzepten wie Informationslogistik, Business Intelligence und Data Warehousing eine unverändert hohe Bedeutung zu [vgl. z. B. So07]. Einher mit der Relevanz des Themas geht eine Erwartungshaltung, die nicht immer erfüllt werden kann. So klagten einer aktuellen Studie von A.T. Kearney zufolge europäische Topmanager über eine mangelhafte Unterstützung der Unternehmensziele durch die IT [AT07]. Insbesondere wachstumsrelevante IT-Bereiche wie Business Intelligence und Customer Relationship Management stehen in der Kritik des Business. Vor allem inkonsistente Daten, zu lange Reaktionszeiten der IT bei der Erfüllung von Anfragen sowie fehlende Transparenz über den konkreten IT-Nutzen werden als kritisch angesehen. Eine Möglichkeit, den zuvor genannten Kritikpunkten entgegen zu steuern, ist die Etablierung einer Governance für die Informationslogistik.

Die Organisation der Unternehmensleitung und -kontrolle zur Sicherung eines optimalen Interessenausgleichs zwischen allen Anspruchsgruppen (Stakeholdern) wird als Corporate Governance bezeichnet [Wi00b]. Abgeleitet von der Corporate Governance beinhaltet IT-Governance Grundsätze, Verfahren und Massnahmen, die sicherstellen, dass mit Hilfe der eingesetzten IT die Geschäftsziele abgedeckt, Ressourcen verantwortungsvoll eingesetzt und Risiken angemessen überwacht werden [MZK03]. Weill und Ross vertreten die Auffassung, dass effektive IT-Governance der wichtigste Einflussfaktor auf den durch IT generierten Nutzen ist [We04; WR04a].

Effektive Governance der Informationslogistik kann also dazu beitragen, dass das durch Informationslogistik erhöhte Nutzenpotenzial realisiert wird. Es ist daher erforderlich zu verstehen, wie Unternehmen eine effektive Informationslogistik-Governance etablieren können, um so bestmöglich zur Erfüllung der Unternehmensziele beizutragen.

Dieser Beitrag beschreibt ein Modell der Informationslogistik-Governance, das als wichtige Komponente eines ganzheitlichen Ansatzes der analytischen Informationsversorgung verstanden werden kann. Als konzeptionelle Grundlagen werden im folgenden Kapitel 2 der St. Galler Ansatz der Informationslogistik und etablierte Definitionen aus der Corporate und IT-Governance vorgestellt. Darauf aufbauend werden in Kapitel 3 eine Definition der Informationslogistik-Governance abgeleitet, die Notwendigkeit einer eigenen Informationslogistik-Governance dargestellt und ein Modell der Informationslogistik-Governance präsentiert, das als Grundlage für die Etablierung einer erfolgreichen Informationslogistik dienen kann. Am Beispiel der E.ON AG wird in Kapitel 4 aufgezeigt, wie die Informationslogistik-Governance im konkreten Fall umgesetzt wurde. In Kapitel 5 werden die Ergebnisse abschließend zusammengefasst und ein Ausblick auf weitere Entwicklungsmöglichkeiten des vorgestellten Modells gegeben.

2 Grundlagen der IL-Governance

2.1 Informationslogistik

Unter Informationslogistik (IL) wird die „Planung, Steuerung, Durchführung und Kontrolle der Gesamtheit der Datenflüsse verstanden, die über eine Betrachtungseinheit

hinausgehen, sowie die Speicherung und Aufbereitung dieser Daten. Dabei werden nur solche Datenflüsse zur IL gezählt, die der Unterstützung von Entscheidungen dienen“ [Wi08]. Als Betrachtungseinheiten können Organisationseinheiten beliebiger Größe, von der Stelle bis hin zum Unternehmensnetzwerk, betrachtet werden. Die IL befasst sich dabei nicht nur mit der Unterstützung von strategischen Entscheidungen, sondern auch mit Managementkontrolle auf der taktischen Ebene und mit operativer Kontrolle sowie mit der Unterstützung strukturierter ebenso wie unstrukturierter Entscheidungen. Neben den dispositiven Prozessen werden zunehmend auch operative Prozesse von der IL mit Informationen versorgt. Damit hat die IL einen breiteren Fokus als das Data Warehousing, das sich zumindest in den älteren Veröffentlichungen nur mit Managementunterstützung befasst [In02]. Während das konstituierende Element des Data Warehousing eine zentrale, integrierte Datenbasis ist [KMU06], abstrahiert die IL weitgehend von der technischen Infrastruktur und fokussiert auf die Datenflüsse. Zudem lag der Schwerpunkt insbesondere der frühen Data Warehouse (DWH)-Literatur auf der Unterstützung des Managements [In02], wobei sich dieses Begriffsverständnis in den letzten Jahren erheblich erweitert und dem der Informationslogistik angenähert hat [KMU06].

Das Konzept der Business Intelligence (BI) ist im Gegensatz zum Data Warehousing weniger klar abgrenzbar. Während das ursprüngliche Verständnis von BI die analytischen Informationssysteme betraf, die auf das DWH zugreifen [LL06; SW02], finden sich in jüngerer Zeit zunehmend umfassendere Interpretationen des Begriffs, die dem hier verfolgten Verständnis von IL ähneln, ohne jedoch den Schwerpunkt auf die Datenflüsse zu legen [GG00; GI01].

2.2 Corporate Governance und IT-Governance

Die IT-Governance wird im Allgemeinen als IT-bezogene Untermenge der Corporate Governance betrachtet [IT03; MZK03]. Corporate Governance bezeichnet die Organisation der Unternehmensleitung und -kontrolle zur Sicherung eines optimalen Interessensausgleichs zwischen allen Anspruchsgruppen (Stakeholdern) des Unternehmens, insbesondere zwischen dem Management und den Aktionären mit Ziel einer langfristigen Wertschöpfung [MZK03; OE04; Wi00b].

Während das Verständnis von Corporate Governance relativ ausgereift und in nationalen und internationalen Standards [OE04; RD04] gefestigt worden ist, bestehen bezüglich des Verständnisses von IT-Governance unterschiedliche Auffassungen [BG05; WPR06]. Gemein haben die meisten Ansätze, dass dem Alignment von Business und IT ein hoher Stellenwert eingeräumt wird [WPR06]. Darunter wird verstanden, dass die IT die Erreichung der Geschäftsziele effektiv und effizient unterstützen soll [HV93].

Ein in der Praxis verbreitetes Verständnis von IT-Governance stammt aus dem Umfeld des IT Governance Institute (ITGI), das mit CobiT den de facto-Standard für IT-Governance Frameworks herausgibt. Das ITGI versteht IT-Governance als „leadership and organisational structures and processes that ensure that the organisation’s IT sustains and extends the organisation’s strategies and objectives“ [IT03]. Dabei liegt der Schwerpunkt auf den Prozessen und Kontrollzielen, was aus dem Hintergrund von CobiT aus dem Bereich der Prüfung von IT-Systemen erklärbar ist.

Diesem Beitrag liegt im Folgenden das weniger auf die Kontrolle fokussierte IT-Governance-Verständnis von Weill und Ross zu Grunde, das anders als CobiT Kontrollziele und Kennzahlen nicht als Teil von IT-Governance sieht. Dieses Verständnis führt die verschiedenen in der Forschung bisher anzutreffenden Ansätze zusammen [BG05] und findet sich auch in dem im vierten Kapitel dargestellten Praxisbeispiel. Weill und Ross verstehen unter IT-Governance „specifying the decision rights and accountability framework for encouraging desirable behaviors in the use of IT“ [WR04a]. IT-Governance ist also ein System von Entscheidungsrechten und Verantwortlichkeiten, das festlegt, wer welche grundlegenden Entscheidungen bezüglich der IT trifft, wer an den Entscheidungsprozessen mitwirkt und wie die beteiligten Stakeholder für die Ausführung ihrer Rollen zur Verantwortung gezogen werden [We04; WR04b]. Sie umfasst damit die Aspekte der Aufbau- und Ablauforganisation, die mit dem Treffen und Durchsetzen der grundlegenden IT-Entscheidungen betraut sind. In dem Aspekt des „erstrebenswerten Verhaltens“ zeigt sich auch in dieser Definition die Bedeutung des Business-IT-Alignments als Treiber der IT-Governance [WR04a].

Die Entscheidungen, die im Rahmen der IT-Governance getroffen werden müssen, betreffen fünf Aspekte [WR04a; WR04b]:

- (abstrakte) Prinzipien für die IT, die in grobgranularer Form umreißen, auf welche Weise die IT das Business unterstützen soll und die die Grundregeln für die IT im Unternehmen festlegen,
- Architekturentscheidungen, die Richtlinien für den logischen und technischen Aufbau der Systemarchitektur und die Integration der Teilsysteme beschreiben,
- Infrastrukturentscheidungen, die die gemeinsam zu nutzenden, zentralen IT-Services definieren,
- Geschäftsanforderungen, die die Anforderungen der Fachbereiche für IT-Anwendungen festhalten und
- Investmententscheidungen, die den Umfang, die Schwerpunkte und die Priorisierungsregeln für IT-Projekte festlegen.

Diese Entscheidungen können auf verschiedene Art und Weise getroffen werden: entweder zentral oder dezentral bzw. unter Führung der Fachbereiche oder unter Führung der IT. Dabei können in einem Unternehmen durchaus unterschiedliche dieser Entscheidungsarchetypen für die jeweiligen Entscheidungen anzutreffen sein [We04]. Die Verantwortung für die IT-Governance liegt aber in jedem Fall beim oberen Management [IT03; WR04a].

Die tatsächlichen Entscheidungen und die daraus hervorgehenden Richtlinien sind dabei nicht Gegenstand der IT-Governance, diese sind Aufgabe des Managements. Im Rahmen von Managementprozessen werden zahlreiche Entscheidungen gefällt – für die oben aufgeführten Schlüsselentscheidungen legt die Governance fest, von wem und wie diese Entscheidungen gefällt werden [WR04a].

In diesem Verständnis folgt die IT-Governance der Geschäftsstrategie. Das Topmanagement legt mit der Strategie und weiteren Leitlinien, wie einer Mission, Wertesystemen

etc., die erstrebenswerten Verhaltensweisen von Teilen des Unternehmens fest. Mittels Vermögensgegenständen wie finanziellen und physischen Vermögenswerten, Mitarbeitern, geistigem Eigentum, Beziehungen und IT werden die Strategie und die erstrebenswerten Verhaltensweisen in der Geschäftstätigkeit umgesetzt. Diese Umsetzung steuert im Bereich der IT die IT-Governance [WR04a].

3 Governance der Informationslogistik

Wie in Abschnitt 2.1 ausgeführt wurde, adressiert Informationslogistik als ganzheitlicher Ansatz die unternehmensweite oder gar -übergreifende Bereitstellung entscheidungsrelevanter Informationen. Zur Schaffung von Synergien müssen Partikularsichten und -interessen unterschiedlicher Organisationseinheiten (oder Funktionalbereiche) im Unternehmen überwunden und ggf. aus historisch gewachsenen, intransparenten und inhomogenen Insellösungen eine Gesamtsicht mit Hilfe einer allgemein getragenen und genutzten Plattform für analytische Zwecke entwickelt werden [DW08b]. Um dieses Ziel zu erreichen, gleichzeitig aber dafür Sorge zu tragen, dass die Informationslogistik die Belange der Fachbereiche optimal unterstützt, werden adäquate und gelebte Governance-Strukturen benötigt. Auch erfordern organisatorische Veränderungen in den Unternehmen (etwa in Folge von Mergers & Acquisitions) zentrale Steuerungs- und Kontrollfunktionen zur Schaffung oder Wahrung von Business Value. Schließlich helfen Governance-Strukturen in der Informationslogistik, die verschiedenen und sich ggf. auch widersprechenden Anforderungen der Fachbereiche richtig zu adressieren und zu priorisieren, um eine gesamthaft möglichst optimale Abdeckung der fachlichen Anforderungen zu erzielen.

Ein weiterer Indikator für die Relevanz adäquater organisatorischer Strukturen (und damit von Governance) in IL-(bzw. DWH)-Projekten zeigt sich in Untersuchungen wie etwa [Wa99], die bestätigen, dass weniger technische als vielmehr organisatorische Gründe zum Scheitern solcher Projekte führen.

3.1 Begriffsverständnis

Im Gegensatz zur IT-Governance, zu der zahlreiche Arbeiten, sei es wissenschaftlicher Natur oder in Praxisbeiträgen, existieren, gibt es bisher kaum Beiträge, die sich gezielt mit IL-Governance beschäftigen. Unter den Begriffen BI-Governance oder DWH-Governance gibt es vereinzelte Arbeiten (wie [ML04; SJG03; WFA04]), weitere sind im Praxisumfeld anzutreffen (z. B. [Gu06]). Eher finden sich noch Arbeiten zu „Data Governance“, die jedoch, wie der Name suggeriert, stark auf Daten bzw. insbesondere Datenqualitätsaspekte fokussieren und damit für den Kontext Informationslogistik zu kurz greifen.

Watson et al. betrachten Data Warehousing Governance als eine Untermenge der IT-Governance [WFA04]. Sie erfordert passende Organisationsstrukturen und Prozesse mit koordinierten Beiträgen aller Stakeholdergruppen und muss die Bereiche Infrastruktur, Nutzung und Projektmanagement abdecken. Gutierrez hingegen unterscheidet drei Per-

spektiven, über die BI-Governance definiert werden kann: Als Priorisierungsmechanismus für BI-Projekte, als eine Sammlung von Richtlinien, Regeln bzw. Empfehlungen und schließlich als die Definition von Rollen und Verantwortlichkeiten für sowohl IT- als auch Business-Stakeholder [Gu06]. Diese drei Aspekte spiegeln das gängige Begriffsverständnis aber auch das Begriffswirrwarr um BI-Governance wider. Die im Folgenden getroffene Definition (bzw. die von Weill und Ross zur IT-Governance) entspricht am ehesten dem dritten Aspekt.

In Anlehnung an das in Abschnitt 2.2 genannte Begriffsverständnis von IT-Governance verstehen wir unter Governance der Informationslogistik (IL-Governance) die aufbau- und ablauforganisatorischen Strukturen, die für das Treffen und Durchsetzen der grundlegenden Entscheidungen in der Informationslogistik verantwortlich sind. Diese Entscheidungs- und Verantwortungsregelung soll das „erstrebenswerte“, also zielkonforme Verhalten in der Nutzung der Informationslogistik fördern. Dabei gilt es im Rahmen der Aufbauorganisation die Organisationsstrukturen, Rollen und Verantwortlichkeiten festzulegen. In der Ablauforganisation wiederum werden Entscheidungs-, Kommunikations- und Überwachungsprozesse definiert [WR04b].

3.2 Die Ausgestaltung der IL-Governance

Die Etablierung einer IL-Governance erfordert die Festlegung entsprechender aufbau- und ablauforganisatorischer Strukturen. Deren Gestaltungsoptionen werden im folgenden Abschnitt 3.2.1 diskutiert. Im Rahmen dieser Strukturen werden dann grundlegende Entscheidungen hinsichtlich bestimmter Aspekte getroffen, die in Abschnitt 3.2.2 in Anlehnung an [We04; WR04a] vorgestellt werden.

Aufbau- und Ablauforganisation

Die aufbau- und ablauforganisatorischen Strukturen im Rahmen der IL-Governance werden anhand folgender Schritte gestaltet: Festlegung der Organisationseinheiten für die IL-Governance, ihre Verankerung in der Aufbauorganisation des Unternehmens, die Definition von Rollen und deren Zuordnung zu den Organisationseinheiten und schließlich entsprechende Entscheidungs-, Kommunikations- und Überwachungsprozesse.

Festlegung der Organisationseinheiten

Hier gilt es zunächst, die grundlegende Entscheidung zu treffen, ob eine (oder mehrere) eigenständige (neue) Organisationseinheiten für die IL-Governance geschaffen oder ob sie in bestehenden Organisationseinheiten verankert werden. Letzteres kann in IL-Einheiten geschehen oder auch in Organisationseinheiten, die für die IT-Governance bereits eingerichtet wurden. Typische Organisationseinheiten sind Komitees (als übergreifende, beispielsweise dem Board unterstellte Steuerungsgremien) und/oder sogenannte Competence Center bzw. Center of Excellence für die Informationslogistik (bzw. für Business Intelligence oder Data Warehousing – je nach Terminologie im Unternehmen) (vgl. [BGH06; Ze06]). In solchen Organisationseinheiten sollten nach Gutierrez alle wichtigen Stakeholdergruppen vertreten sein, also IT- und Fach-Vertreter sowie Repräsentanten aus dem CFO-Bereich [Gu06].

Verankerung der Organisationseinheiten in der Aufbauorganisation des Unternehmens

Werden für die IL-Governance neue Organisationseinheiten geschaffen, muss festgelegt werden, wie diese in die bestehende Aufbauorganisation des Unternehmens (häufig funktional, regional oder nach Sparten angeordnet) eingepasst werden. Das Praxisbeispiel in Abschnitt 4 zeigt, dass auch orthogonal zu bestehenden Strukturen neue Anordnungskriterien zum Zuge kommen können – in diesem Fall werden die sogenannten Center of Excellence entlang der Geschäftsprozesse positioniert.

Des Weiteren stellt sich die Frage, ob die Organisationseinheiten zur IL-Governance auf IT- oder Fach-Seite angesiedelt werden. In der unternehmerischen Praxis finden sich hier unterschiedliche Ausprägungen, wenn sie etwa dem CIO, dem CFO oder dem Leiter eines Geschäftsbereichs unterstellt werden (vgl. [Ho08]). Die Frage lässt sich nicht allgemeingültig beantworten, da auch hier die Unternehmensspezifika in Betracht gezogen werden müssen; etwa inwieweit die IT bereits strategisch ausgerichtet ist und inwiefern sie fachlich getrieben operiert. Virtuelle Organisationseinheiten, die in einer Matrixstruktur parallel zur bestehenden Organisation aufgebaut werden, lassen sich in der Regel schneller, einfacher und kostengünstiger etablieren, auf der anderen Seite besteht hier stets die Gefahr, dass die beteiligten Mitarbeiter sich bei auftretenden Zielkonflikten im Zweifelsfall den Interessen des Managements, dem sie primär unterstehen, unterordnen. Kommunikations- und Abstimmungsprozesse sind in virtuellen organisatorischen Einheiten ebenfalls erschwert.

In [We04; WR04a] (vgl. auch Abschnitt 2.2) und in [SJG03] werden jeweils Grundformen der IT-Governance vorgestellt, die sich im Wesentlichen darin unterscheiden, welchen Einflussgrad jeweils IT- bzw. Fachanwender haben. Sie lassen sich prinzipiell auch auf die IL-Governance übertragen.

Definition von Rollen und Zuordnung zu Organisationseinheiten

Die Organisationseinheiten für die IL-Governance benötigen eine adäquate Besetzung durch Vertreter der verschiedenen Stakeholdergruppen sowie einen Mix an Know How-Trägern zur Abdeckung aller grundlegenden Entscheidungsfragen (vgl. Abschnitt 3.2.2). In [BGH06] wird gezeigt, welche Kompetenzen in einem IL-/BI-Competence Center benötigt werden.

Tabelle 1 zeigt beispielhaft, welche Rollen unterschieden werden könnten. Die genaue Ausgestaltung der Rollen hängt dann jeweils von den Unternehmensspezifika ab. Auch gilt es zu berücksichtigen, welche Rollen bereits in der IT-Governance besetzt sind – in solchen Fällen bietet es sich ggf. an, Rollen bei bestimmten Personen zusammenfassen, so dass beispielsweise ein Data Steward für die Datenqualität sowohl in operativen als auch in dispositiven Systemen verantwortlich ist. Auch bei dem in Abschnitt 4 geschilderten Praxisbeispiel werden Rollen zusammengefasst.

Rolle	Aufgaben
Architekt	Entwicklung und Überwachung der logischen Architektur und Definition technischer Standards
Projektportfolio-Manager	Planung, Koordination und Überwachung von Projekten
Projekt-Manager	Führung von Projekten und Sicherstellung, dass sie den verpflichtenden Architekturstandards genügen und gleichzeitig die fachlichen Anforderungen erfüllen
Data Steward	Sicherstellung der Datenqualität, Korrektheit der Transformationslogik und Lieferung von Daten an Data Warehouse(s) oder Data Marts
Anforderungs-Manager	Sammlung, Konsolidierung und Priorisierung von Anforderungen als Vorbereitung für die Projektplanung
Anwender-Vertreter	Repräsentation der Anwender und ihrer Interessen in Projekten, Projektportfolioplanung und Architekturdesign
Application-Manager	Sicherstellung der funktionalen Eignung und Konsistenz von Applikationen
IT-Spezialist	Entwicklung und Wartung technischer Lösungen (DWH und zugehörige Applikationen bzw. Schnittstellen zu den Quellsystemen)
Qualitätsmanager	Überwachung von Anforderungen, Projekten, Applikationen und Prozessen und Prüfung möglicher Verbesserungsmaßnahmen
Security Manager	Sicherstellung der Datensicherheit (autorisierte Datenzugriffe)

Tabelle 1: Rollen in der IL-Governance

Ergänzend werden den Rollen noch Verantwortlichkeiten und Zuständigkeiten zugewiesen, beispielsweise in Form einer so genannten RACI-Matrix. Dann wird für jede Aktivität bestimmt, ob die Rolle (Person) diesbezüglich verantwortlich in der Durchführung (responsible), verantwortlich im Sinne der Genehmigung (accountable), befragt (consulted) oder informiert (informed) wird bzw. ist.

Entscheidungs-, Kommunikations- und Überwachungsprozesse

IL-Governance erfordert neben den aufbauorganisatorischen Strukturen gleichermaßen adäquate Prozesse. Das CobiT-Referenzmodell [IT05] bietet hierfür einen Leitfaden zur Durchführung der Kontroll- und Führungsaufgaben im Rahmen der IT-Governance an. Es stellt ein Modell von generell anwendbaren Kontrollzielen bereit, die in einem Unter-

nehmen umgesetzt werden sollten, um eine verlässliche Anwendung der IT zu gewährleisten. CobiT identifiziert zudem 34 kritische Prozesse innerhalb der IT, welche für ein angemessenes Management der IT erfolgsbestimmend sind. Auch wenn die Kontrollziele und Prozesse in CobiT einen allgemeingültigen Anspruch für die gesamte IT haben, so sind sie doch in großen Teilen auf die Informationslogistik übertragbar und können als Leitfaden zur Definition und Etablierung der Entscheidungs- Kommunikations- und Überwachungsprozesse in der IL-Governance in Einklang mit den vorgesehenen Rollen und Verantwortlichkeiten (s. oben) herangezogen werden.

Gestaltungsobjekte der IL-Governance

Im Rahmen der IL-Governance werden grundlegende Entscheidungen getroffen und Richtlinien (oft „Principles und Guidelines“ genannt) aufgestellt. Für die IT-Governance identifizieren Weill und Ross fünf Gestaltungsobjekte, hinsichtlich derer diese Entscheidungen getroffen werden [WR04a; WR04b] (vgl. auch Abschnitt 2.2). Sie lassen sich im Wesentlichen auch auf die Informationslogistik übertragen.

- *IL-Prinzipien*: Ausgehend von den Unternehmenszielen und Geschäftsprinzipien sollten Prinzipien in der Informationslogistik festgelegt werden, die deren optimale Unterstützung zum Ziele haben. Sie spezifizieren das „wünschenswerte“ Verhalten der Informationslogistik. Beispiel eines solchen IL-Prinzips könnte etwa die Forderung sein, dass alle Informationen für Managemententscheidungen zeitnah zur Verfügung gestellt werden.
- *Architektur*: In diesem Kontext sind grundlegende Architekturentscheidungen zu treffen, wie beispielsweise den Grad der Zentralisierung oder Aufbau der (Mehr-)Schichtenarchitektur (Materialisierung von DWH und/oder Data Marts, unabhängige versus abhängige Data Marts, etc.) festzulegen. Des Weiteren, ob und welchem Architekturparadigma (wie etwa serviceorientierten Architekturen) gefolgt wird, bzw. ob und wie Rückschreiben in das Data Warehouse erlaubt wird. Aber auch Aussagen zu projektübergeordneten Themen wie Datenqualitäts-, Stamm- und Metadatenmanagement sollten formuliert werden, etwa z. B. „Alle Daten, die in das Data Warehouse geladen werden, sind zuvor einer Qualitätskontrolle zu unterziehen.“
- *Infrastruktur*: Hierunter können nicht nur Aussagen zur Infrastruktur im eigentlichen Sinne, also Hardware- und Softwarelandschaft in der Informationslogistik (bestimmte Hersteller- oder Toolpräferenzen oder Buy- versus Make-Festlegung) gezählt werden, sondern beispielsweise auch Aussagen zu Fertigungstiefe und Sourcing (welche Services bzw. Leistungen die Informationslogistik selber erbringt oder fremdvergift). Grundlegende Entscheidungen zur Kosten- und Leistungsverrechnung in der Informationslogistik sind hier ebenfalls zu treffen.
- *Geschäftsanforderungen*: Übertragen auf den Kontext der Informationslogistik muss z. B. festgelegt werden, mit welchem Angebot an Produkten der Informationslogistik das Business optimal unterstützt werden kann und welche neuen Möglichkeiten für das Unternehmen (Produkte, Geschäftsmodelle und -prozesse, etc.) sich daraus ergeben können.

- *Investment und Priorisierung*: Wiederholt wird betont (wie in [ML04; WFA04]), dass Anforderungen aus dem Business und Projektvorhaben eine richtige Priorisierung erfahren müssen, um sicherzustellen, dass die Bedürfnisse des Business optimal erfüllt werden können. Hier kommt den Organisationseinheiten der IL-Governance insbesondere noch die Verantwortung zu, für optimale Synergien zwischen einzelnen Bereichen zu sorgen, Partikularinteressen zu überwinden und Redundanzen zu vermeiden.

Ein Abgleich dieser Gestaltungsobjekte mit den in [WFA04] genannten Gebieten¹, die eine DWH-Governance abdecken muss, zeigt, dass diese in den oben genannten enthalten sind.

Im folgenden Abschnitt wird anhand eines Praxisbeispiels gezeigt, wie eine IL-Governance konkret aufgebaut und umgesetzt werden kann.

4 Informationslogistik-Governance bei der E.ON AG als Teil eines ganzheitlichen Ansatzes

E.ON ist der weltweit größte private Energiedienstleister. Für derzeit rund 40 Millionen Kunden (Privat- und Geschäftskunden, Industriebetriebe sowie regionale und lokale Energieunternehmen) in über 20 Ländern stellt E.ON Lösungen für Strom und Gas zur Verfügung. Dabei konzentriert sich das Unternehmen auf die Zielmärkte Zentraleuropa, Großbritannien, Nordeuropa und den Mittleren Westen der USA. Nach einer Transformationsphase hin zu einem fokussierten Energieunternehmen (2000-2003) und einer Integrations- und Investitionsphase (2003-2005) befindet sich der E.ON Konzern seit 2005 in einer Wachstumsphase, die eine aktive, zentrale Steuerung des Konzerns durch das Corporate Center und eine fortwährende Integration aus Sicht der Strategie und der Organisation erfordert. Dies führt zu einer Integration und Zentralisierung von Funktionen, die organisationseinheitenübergreifende Datenflüsse (in Abbildung 1 rechte Seite) erforderlich macht. Die Datenflüsse sind nicht mehr nur innerhalb der Organisationseinheiten anzutreffen (linke Seite in der Abbildung). Insbesondere anhand der sich verändernden fachlichen und organisatorischen Rahmenbedingungen sowie der damit einhergehenden Anforderungen zur Informationsversorgung wurde deutlich, dass eine strukturierte, ganzheitliche Gestaltung der Informationslogistik unumgänglich ist. Dieses Vorhaben wird in Form des Framework Information Logistics @ E.ON, kurz FIL.ON, derzeit konzernweit erarbeitet und etabliert. Einen Teilaspekt der ganzheitlichen Gestaltung der Informationslogistik bildet die IL-Governance, die im Folgenden, nach einer überblicksartigen Schilderung des FIL.ON, detailliert erläutert wird.

¹ Identifikation von Business Needs, Alignment mit den Unternehmenszielen, Ermittlung der Informationsbedürfnisse, Priorisierung, Etablierung von Datendefinitionen und -modellen und Sicherstellung der Datenqualität

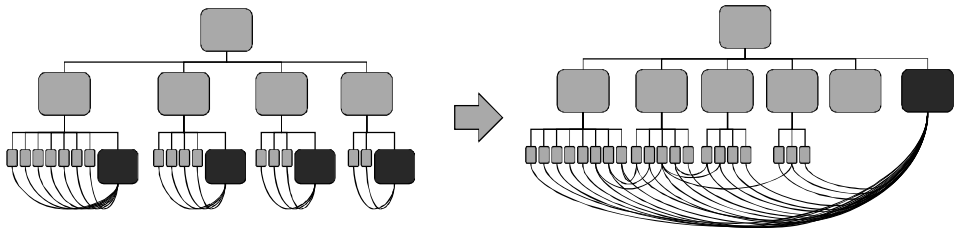


Abbildung 1: Aus Zentralisierung und Integration von Funktionen resultieren betrachtungseinheitenübergreifende Informationsbedürfnisse

4.1 Der ganzheitliche Ansatz einer IL-Strategie: Das Framework Information Logistics @ E.ON

Informationslogistik definiert bei E.ON die Grundlage, um Informationen zur richtigen Zeit mit der notwendigen Qualität bereitzustellen. Dabei werden alle Datenflüsse betrachtet, die über eine Organisationseinheit hinaus entscheidungsrelevant sind. Diese Datenflüsse werden so aufgesetzt und vernetzt, dass wichtige Geschäftsinformationen generiert und geteilt werden können. Auch die Speicherung und weitere Aufbereitung der Daten ist im Fokus der IL-Strategie. Im E.ON-Konzern gab es bisher keine übergreifende IL-Strategie. Die neue Organisationsstruktur des E.ON-Konzerns erfordert eine verbesserte Informationsgrundlage für Entscheider. Zur erfolgreichen Steuerung der Unternehmensentwicklung ist das Management von Informationsflüssen über die Grenzen von Business und Market Units hinweg nötig.

Auf Basis dieses neuen Ansatzes entwickelt das Projekt FIL.ON einen konkreten konzernweit gültigen und akzeptierten Rahmen für alle IL-Projekte im Konzern. Das FIL.ON-Team erarbeitet unter anderem gemeinsame Methoden, Vorlagen, Strukturen und IL-Architekturprinzipien. So können komplexe Daten aus unterschiedlichen Prozessen künftig schnell und effizient verknüpft werden. Das Ergebnis ist eine verbesserte Informationsgrundlage für Entscheider.

Ziele des FIL.ON-Projekts und der daraus resultierenden IL-Strategie sind u.a. eine verstärkte Ausrichtung der IL anhand der Erfordernisse des Business, die Bereitstellung von Grundlagen für eine flexible IL-Architektur sowie die Harmonisierung und Standardisierung der IL zur Erhöhung der Transparenz und Effizienz.

Der fachliche und organisatorische Teil des Frameworks beinhaltet u. a. eine Methode zur Informationsbedarfsanalyse und ein globales KPI-Repository zur einheitlichen Definition von Geschäftsmetriken. Im Bereich der Datenqualität werden zunächst eine eindeutige Definition von Datenqualität etabliert und Datenqualitätslevel festgelegt, die die Relevanz der Daten in sämtlichen Nutzungsprozessen beinhalten. Zusätzlich wird ein Prozessmodell zur Verbesserung der Datenqualität bereitgestellt. Im Rahmen einer Information-Life-Cycle-Management-Strategie, die Bestandteil des Frameworks ist, wird eine einheitliche Definition eines Information Life Cycle aus fachlicher Sicht etabliert. Das Prozessmodell zur gruppenweiten Klassifizierung des Life Cycles ist ebenfalls enthalten. Beide zuvor genannten Themen sind Bestandteil der übergreifenden Information

Governance und des Rollenmodells, in dem verschiedene Arten von Verantwortlichkeiten und Aufgaben definiert sind. Abgerundet wird der Aufbau einer fachlichen IL-Methodik durch ein konzernweit einheitliches Terminologie-Management für Fachbegriffe aus dem Bereich der analytischen Applikationen, so dass insbesondere ein einheitliches Verständnis dieser Begriffe zwischen Fach- und IT-Abteilungen geschaffen werden kann.

Der informationstechnische Teil des Frameworks enthält die Grundsätze der IL-Landschaft, Architekturrichtlinien, Verrechnungsmodelle für analytische Applikationen und Konzepte für die Evolution der IL-Architektur. Anhand von Best-Practice-Beispielen ähnlich strukturierter Unternehmen wurden beispielsweise Alternativen für zukünftige Zielarchitekturen identifiziert und aufgrund der E.ON-spezifischen Anforderungen bewertet. Aufbauend auf diesen Erkenntnissen werden Architekturrichtlinien entwickelt, durch deren gruppenweite Einhaltung die Transparenz, Wartbarkeit und Wiederverwendung von Daten und Systemkomponenten, aber auch die Effizienz erhöht werden soll.

In Abbildung 2 werden die Ziele sowie Wegbereiter zur Erreichung der Ziele („Enabler“) bzw. Ergebnisse zusammenfassend dargestellt.

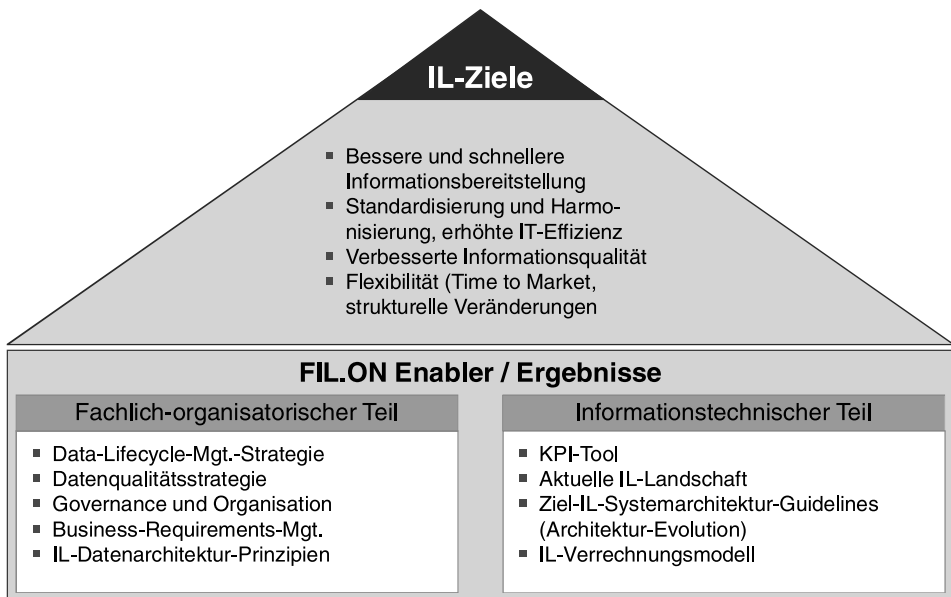


Abbildung 2: Das Framework Information Logistics – der ganzheitliche Ansatz der Informationslogistik der E.ON AG

Im Sinne einer ganzheitlichen, strukturierten Gestaltung der Informationslogistik bildet das Framework die Grundlage aller weiteren Aktivitäten. Auf das Thema Governance der Informationslogistik wird im Folgenden detailliert eingegangen.

4.2 Informationslogistik-Governance bei der E.ON AG

Einen Schwerpunkt des Frameworks bildet der Teilaspekt „IL-Governance und Organisation“. Die IL-Governance wird im E.ON-Konzern als zur Umsetzung gruppenweiter Ziele notwendig erachtet. Das Verständnis von Governance entspricht in weiten Teilen der in Abschnitt 3.1 entwickelten allgemeinen Definition von IL-Governance bzw. dem in Abschnitt 3.2 entwickelten IL-Governance-Modell.

Die Organisationsstruktur (Aufbauorganisation) der Governance besteht aus Rollen und deren Beschreibungen, Zuordnung der Rollen zu Organisationseinheiten, Gremien oder Personen, sowie den Beziehungen zwischen den zugeordneten Rollen. Als wichtigste Rollen der IL gelten der Business-Information-Manager, der Information-Architekt und der IL-Architekt. In Abbildung 3 werden diese Rollen detailliert beschrieben.




 Business-Information-Manager	<ul style="list-style-type: none">• Anforderungsmanagement (nur Bereich IL)• Koordination von Projekten
 Information-Architekt	<ul style="list-style-type: none">• Kommunikation und IL-Marketing• Vermittlerartige Schnittstelle zu den Fachbereichen• IL-Problem-Management• Prozess-Design und IL-Anforderungsaufnahme
 IL-Architekt	<ul style="list-style-type: none">• Entwicklung und Monitoring der logischen Architektur• Applikations-Design (entsprechend FIL.ON Guidelines)• Operative Verantwortung für die Compliance mit den Informations-Architektur-Guidelines• Ausarbeitung der Informations-Architektur in Compliance mit den IL-Architektur-Guidelines• Überprüfung von Projekten, Geschäftsprozessen und Anwendungen in Bezug auf Compliance mit Prinzipien und Guidelines auf Ebene des Governance-Boards

Abbildung 3: Die drei wichtigsten IL-Rollen und deren Beschreibung

Eine Kernfrage der Governance ist, wie die Rollen den Organisationseinheiten des E.ON-Konzerns, also dem Corporate Center, den Markt- (sog. Market Units) und den Geschäftseinheiten (sog. Business Units) zugeordnet werden können. Um die IL-Expertise in fachlichen Themen zu bündeln und einheitenübergreifende Synergiepotenziale zu realisieren, wurde entschieden, die IL-Governance rein funktional zu verankern. Für die IL-Governance existieren drei Arten von Organisationseinheiten: Das Governance-Board besitzt eine reine Steuerungsfunktion. Typische Aufgaben dieses Gremiums sind die Koordination von IL-Governance als Teil der Corporate Governance, die Verabschiedung der IL-Strategie und die Management-Unterstützung der Durchsetzung der Richtlinien und Guidelines. Die funktionale Verankerung der IL-Governance erfolgt durch so genannte Centers of Excellence (CoE). Für die Erstellung und Umsetzung konzernweiter IL-Belange existiert ein CoE des Corporate Centers, kurz CoE-CC. Das CoE-CC erarbeitet gruppenweit gültige Architekturprinzipien, Richtlinien und Standards, die dem Governance-Board als Entscheidungsgrundlage vorgelegt werden.

Für die Kerngeschäftsprozesse existieren weitere Centers of Excellence, die als virtuelle CoEs (vCoEs) bezeichnet werden. Diese Centers weisen fachliche Expertise ihrer jewei-

ligen Kerngeschäftsprozesse auf und stellen für die zu diesen Geschäftsprozessen zugehörigen Fachanwender den primären Ansprechpartner für IL-Belange dar. Werden aus einem Geschäftsprozess heraus Anforderungen definiert, die auch andere Geschäftsprozesse betreffen, so wird aufgrund des konzernweiten Interesses einer einheitlichen Umsetzung das vCoE zum Koordinator, in dessen Bereich die Anforderung gestellt wurde. Somit wird eine Geschäftsprozess übergreifende Umsetzung gewährleistet. In Abbildung 4 wird die Aufbauorganisation, bestehend aus dem CC-CoE, den vCoEs und der Zuordnung der CoEs zu den Kernprozessen, dargestellt.

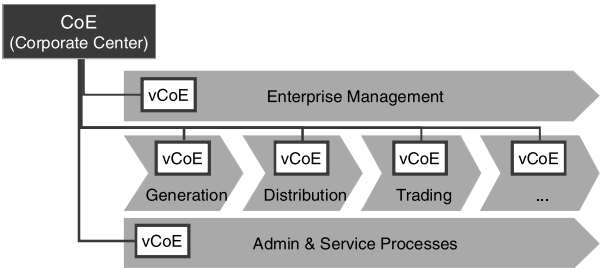


Abbildung 4: Rein funktionale Ausrichtung der IL-Aufbauorganisation: (Virtuelle) Centers of Excellence

Zur Ablauforganisation der Informationslogistik bei E.ON werden Prozesse wie z. B. das Anforderungsmanagement und das IL-Portfolio-Management gezählt. Mit einem dedizierten IL-Anforderungsmanagement soll erreicht werden, dass redundante Anforderungen vermieden bzw. konsolidiert, die Wiederverwendung gefördert und letztendlich also Anforderungen im Sinne einer ganzheitlichen Informationslogistik strategiekonform und effizient umgesetzt werden. Abbildung 5 stellt den Anforderungsaufnahmeprozess zusammenfassend dar. Durch einen IL-Anforderer wird eine (neue) Anforderung an das dem Anforderer zugehörige vCoE herangetragen. Die Anforderung wird analysiert, überprüft und dann ggf. zur Umsetzung freigegeben. Die Umsetzung erfolgt dann z. B. in der Projekt-Organisation durch einen IT-Dienstleister.

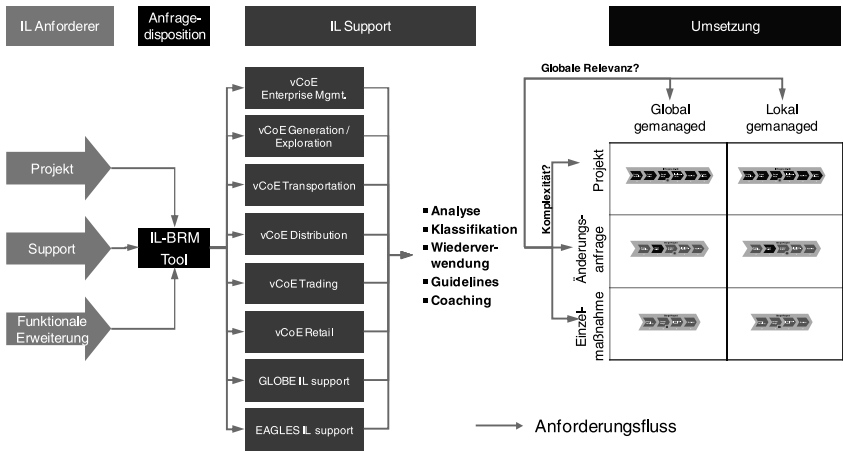


Abbildung 5: High Level-Sicht auf den Anforderungsaufnahmeprozess

Als Schlüssel zum Erfolg wird das richtige Maß an Governance angesehen (vgl. Abbildung 6). Aus Sicht der Konzernführung kann die Governance im Sinne von „Laissez-faire“, „Totale Kontrolle“ oder einem beliebigen Wert zwischen diesen beiden Extrema implementiert werden. Bei einer „soften“ Governance ist die Fähigkeit zur Erfüllung kurzfristiger lokaler Informationsbedürfnisse am ehesten gegeben und nimmt mit zunehmender Governance aufgrund von immer stärker werdender Regulierung ab. Die Gesamtkosten sind hoch, da zur Befriedigung übergreifender Informationsbedürfnisse extreme Aufwände, z. B. bei der Datenkonsolidierung und –integration, anfallen. Das andere Extremum zeichnet sich durch einen sehr hohen Grad an Governance mit entsprechend stark ausgeprägten Organisationsstrukturen, damit verbunden hohen Kosten und der geringen Fähigkeit zur Erfüllung von lokalen und nicht-lokalen Informationsbedürfnisse aufgrund (zu) starker Regulierung aus. Zwischen den beiden Extremwerten lässt sich ein Optimum ausmachen, das sich durch geringe Gesamtkosten, gute Datenintegration und geringen Aufwand zum Aufbau von Organisationsstrukturen auszeichnet. Dieses optimale Maß an Governance gilt es zu erreichen.

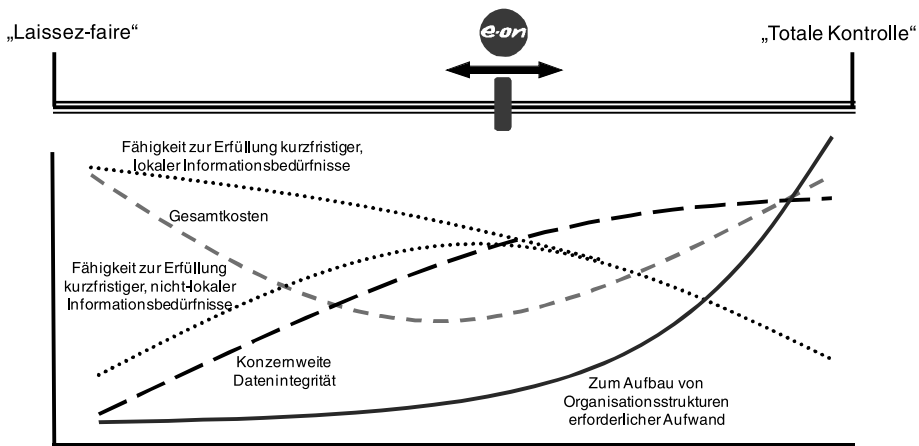


Abbildung 6: Das richtige Maß an Governance

Vergleicht man die bei E.ON etablierten bzw. im Aufbau befindlichen Governance-Strukturen mit dem in Abschnitt 3 entwickelten IL-Governance-Modell, so sind klare Parallelen, wie z. B. die Definition von Organisationsstrukturen in Form der Aufbau- und Ablauforganisation feststellbar. Eine Besonderheit bei E.ON ist jedoch, dass auch die Entscheidungen, welche IL-Projekte durchgeführt (Portfolio-Management) und welche Anforderungen aufgenommen (Business-Requirements-Management) werden, als Teil der Governance angesehen werden. Die IL-Governance im E.ON-Konzern ist also auch mit Managementaufgaben betraut.

5 Zusammenfassung und Ausblick

Die Governance der Informationslogistik ist das Entscheidungsrechte- und Verantwortungssystem, das die Informationslogistik im Sinne der Geschäftsziele steuert. In

aufbauorganisatorischer Hinsicht ergeben sich verschiedene Möglichkeiten, die IL-Governance in bestehenden oder neu zu schaffenden Organisationseinheiten zu verankern, wobei jeweils die Fachbereiche oder die IT die Führungsrolle übernehmen können und unterschiedliche Zentralisierungsgrade in Frage kommen. Ein Rollenkonzept zeigt die im Rahmen der IL-Governance zuzuweisenden Aufgaben. Für die Prozesse der IL-Governance existiert mit CobiT ein Referenzmodell, das auch für die Kontrolle der Umsetzung zahlreiche Hinweise liefert. Das Praxisbeispiel zeigt die Umsetzung der IL-Governance im E.ON-Konzern. Mit virtuellen Kompetenzzentren wurde dort eine der dezentralen Konzernorganisation angemessene Organisationsform etabliert. Prinzipien und Richtlinien regeln die Ausgestaltung der Gestaltungsobjekte.

Für die weitere Forschung und für die Praxis ergeben sich interessante Fragen hinsichtlich der Zusammenhänge von Corporate Governance, IT-Governance und IL-Governance. Insbesondere gilt es die Frage zu beantworten, in wie weit sich die IL-Governance aus den übergeordneten Governance-Systemen ableiten lässt. Ebenfalls interessant ist in diesem Kontext die Frage, ob die IT-Governance und die IL-Governance parallel gestaltet werden sollten oder ob auch divergierende Ausgestaltungen sinnvoll sein können.

6 Literaturverzeichnis

- [AT07] A.T. Kearney: IT-Abteilungen bremsen Unternehmenswachstum. A.T. Kearney, Düsseldorf 2007.
- [BG05] Brown, A. E.; Grant, G. G.: Framing the Frameworks: A Review of IT Governance Research. In: Communications of the Association for Information Systems 15 (2005) May, S. 696-712.
- [BGH06] Burton, B. et al.: Organizational Structure: Business Intelligence and Information Management. Gartner Research G00138940, 2006.
- [De05] De Voe, L.; Neal, K.: When Business Intelligence Equals Business Value. In: Business Intelligence Journal 10 (2005) 3, S. 57-63.
- [DS96] Dean Jr., J. W.; Sharfman, M. P.: Does Decision Process Matter? A Study of Strategic Decision-Making Effectiveness. In: The Academy of Management Journal 39 (1996) 2, S. 368-396.
- [DW08a] Dinter, B.; Winter, R. (Hrsg.): Integrierte Informationslogistik. Springer, Berlin et al. 2008.
- [DW08b] Dinter, B.; Winter, R.: Strategie der Informationslogistik. In: Dinter, B.; Winter, R. (Hrsg.): Integrierte Informationslogistik. Springer, Berlin, Heidelberg 2008, S. 63-82.
- [GG00] Grothe, M.; Gentsch, P.: Business Intelligence. Addison-Wesley, München 2000.
- [GI01] Gluchowski, P.: Business Intelligence – Konzepte, Technologien und Einsatzbereiche. In: HMD – Praxis der Wirtschaftsinformatik 222 (2001), S. 5-15.
- [Gu06] Gutierrez, N.: Business Intelligence (BI) Governance. Infosys Technologies Limited 2006.

- [Ho08] Hostmann, B.: BICC: Shifting the Focus From Technology to Core Business Competency. 2008.
- [HV93] Henderson, J. C.; Venkatraman, N.: Strategic Alignment: Leveraging Information Technology for Transforming Organizations. In: IBM Systems Journal 32 (1993) 1, S. 4-16.
- [In02] Inmon, W.: Building the Data Warehouse. 3. Aufl., Wiley, New York 2002.
- [IT03] IT Governance Institute: Board Briefing on IT Governance, 2nd Ed. IT Governance Institute, Rolling Meadows 2003.
- [IT08] Institute, I. T. G.: CobiT 4.0: Control Objectives, Management Guidelines, Maturity Models. <http://www.isaca.org>, Abruf am 28.03.2008.
- [KD04] Kohli, R.; Devaraj, S.: Realizing the Business Value of Information Technology Investments: An Organizational Process. In: MIS Quarterly Executive 3 (2004) 1, S. 53-68.
- [KMU06] Kemper, H.-G.; Mehanna, W.; Unger, C.: Business Intelligence – Grundlagen und praktische Anwendungen. Eine Einführung in die IT-basierte Managementunterstützung. 2. Aufl., Vieweg 2006.
- [LL06] Laudon, J.; Laudon, K.: Management Information Systems: Managing the Digital Firm. 10. Aufl., Prentice Hall 2006.
- [ML04] Matney, D.; Larson, D.: The Four Components of BI Governance. In: Business Intelligence Journal 9 (2004) 3, S. 29-36.
- [MZK03] Meyer, M.; Zarnekow, R.; Kolbe, L. M.: IT-Governance – Begriff, Status quo und Bedeutung. In: Wirtschaftsinformatik 45 (2003) 4, S. 445-448.
- [OE04] OECD: OECD-Grundsätze der Corporate Governance. OECD 2004.
- [RD04] Regierungskommission Deutscher Corporate Governance Kodex: Deutscher Corporate Governance Kodex. 2007.
- [SJG03] Sujitparapitaya, S.; Janz, B. D.; Gillenson, M.: The Contribution of IT Governance Solutions to the Implementation of Data Warehouse Practice. In: Journal of Database Management 14 (2003) 2, S. 52-69.
- [So07] Sommer, D.: Spending Preferences for Business Intelligence and Information Infrastructure, 2007. Gartner, Stamford 2007.
- [SW02] Strauch, B.; Winter, R.: Stichwort „Business Intelligence“. In: Bellmann, M.; Krcmar, H.; Sommerlatte, T. (Hrsg.): Praxishandbuch Wissensmanagement – Strategien, Methoden, Fallbeispiele. Symposium, Düsseldorf 2002, S. 439-448.
- [Wa99] Watson, H. J. et al.: Data warehousing failures: case studies and findings. In: Journal Of Data Warehousing 4 (1999) 1, S. 44-55.
- [We04] Weill, P.: Don't Just Lead, Govern: How Top-Performing Firms Govern IT. In: MIS Quarterly Executive 3 (2004) 1, S. 1-17.
- [WFA04] Watson, H. J.; Fuller, C.; Ariyachandra, T.: Data warehouse governance: best practices at Blue Cross and Blue Shield of North Carolina. In: Decision Support Systems 38 (2004) 3, S. 435-450.

- [Wi00a] Winter, R.: Zur Positionierung und Weiterentwicklung des Data Warehousing in der betrieblichen Applikationsarchitektur. In: Jung, R.; Winter, R. (Hrsg.): Data Warehousing Strategie. Springer, Berlin et al. 2000, S. 127-139.
- [Wi00b] Witt, P.: Corporate Governance im Wandel. In: Zeitschrift Führung + Organisation 69 (2000) 3, S. 159-163.
- [Wi08] Winter, R. et al.: Das St. Galler Konzept der Informationslogistik. In: Dinter, B.; Winter, R. (Hrsg.): Integrierte Informationslogistik. Springer, Berlin 2008, S. 1-16.
- [WPR06] Webb, P.; Pollard, C.; Ridley, G.: Attempting to Define IT Governance: Wisdom or Folly? In: Proceedings of the 39th Hawaii International Conference on Systems Sciences (HICSS'06), Big Island, Hawaii 2006.
- [WR04a] Weill, P.; Ross, J. W.: IT Governance – How Top Performers Manage IT. Harvard Business School Press, Boston 2004.
- [WR04b] Weill, P.; Ross, J. W.: IT Governance on One Page. Massachusetts Institute of Technology, Center for Information Systems Research (CISR), Cambridge MA 2004.
- [Ze06] Zeid, A.: Your BI Competency Center: A Blueprint for Successful Deployment. In: Business Intelligence Journal 11 (2006) 3, S. 14-20.

BI-Bebauungsplanung im Rahmen der BI-Strategie des Volkswagen-Konzerns

Thorsten Sommer*, Wolfgang Bayer*, Stefan Rosenbaum*, Birgit Wendlandt*,
Alexander Overmeyer**

*Volkswagen AG
ITP-KF Business Intelligence
Brieffach 011/18020
D-38436 Wolfsburg
bi-cc@volkswagen.de

**Business Intelligence Solutions
Steria Mummert Consulting AG
Domagkstraße 34
D-80807 München
alexander.overmeyer@steria-mummert.de

Abstract: Die BI-Landschaft im Volkswagen-Konzern ist durch eine hohe Komplexität und Heterogenität gekennzeichnet. Die Steuerung und Beherrschbarkeit dieser Landschaft ist durch die Vielfalt der eingesetzten Prozesse, Architekturen und Technologien stark beeinträchtigt. Daraus folgt ein erhöhter Aufwand bei der konzernweiten Planung eines gezielten und effizienten Einsatzes von BI und schließlich eine mangelnde Geschäftsunterstützung durch BI.

Dieser Beitrag stellt das Konzept und das Umfeld der BI-Bebauungsplanung zur gezielten Planung und Weiterentwicklung der BI-Landschaft im Rahmen der BI-Strategie des Volkswagen-Konzerns vor. Insbesondere mit dem BI-Bebauungsplan steht dem Konzern ein Werkzeug zur Verfügung, das langfristig zur Umsetzung der Unternehmensziele und der in der BI-Strategie definierten Ziele bezüglich Harmonisierung und Standardisierung, Reduzierung der Systemvielfalt und des Schließens von Prozesslücken dient.

Der BI-Bebauungsplan schafft u. a. durch geeignete Visualisierungsformen die notwendige Transparenz über die fachlichen, technischen und organisatorischen Strukturen der BI-Landschaft und ermöglicht damit die Verwaltung der Komplexität der BI-Landschaft. Dadurch bildet der BI-Bebauungsplan die Grundlage für Entscheidungsfindung und Steuerung der BI-Governance im Rahmen der Ausgestaltung der wohldefinierten BI-Plattform. Die BI-Plattform wird den Konzern in die Lage versetzen, flexibel und schnell auf sich ändernde Marktanforderungen zu reagieren und damit langfristig eine führende Position am Markt zu besetzen.

1 Einleitung

Der Volkswagen-Konzern mit Sitz in Wolfsburg ist einer der führenden Automobilhersteller weltweit und der größte Automobilproduzent Europas mit einem Umsatz von 108,897 Mrd. € und einem Ergebnis nach Steuern von 4,122 Mrd. € im Geschäftsjahr 2007.

Zu den acht Konzernmarken aus sechs europäischen Ländern zählen neben Volkswagen PKW, Audi, SEAT, Skoda und Volkswagen Nutzfahrzeuge ebenfalls Bentley, Bugatti und Lamborghini. Jede Marke hat ihren eigenständigen Charakter und operiert selbstständig im Markt. Dabei reicht das Angebot von verbrauchsoptimalen Kleinwagen bis hin zu Fahrzeugen der Luxusklasse. Im Bereich der Nutzfahrzeuge beginnt das Angebot bei Pick-up-Fahrzeugen und reicht bis zu Bussen und schweren Lastkraftwagen.

Im Jahr 2007 produzierte das Unternehmen über sechs Millionen Fahrzeuge in 48 Fertigungsstätten. Im selben Jahr wurden konzernweit 34 neue Modelle und Produktaufwertungen präsentiert. Seine Fahrzeuge bietet das Unternehmen in mehr als 150 Ländern an. Darüber hinaus gibt es eine Vielzahl fahrzeugbezogener Dienstleistungen – wie Kundendienst und Originalteile sowie kundenbezogene Dienstleistungen wie Finanzierung, Leasing, Vermietung und Versicherungen, deren Erstellung ohne den Einsatz von IT überhaupt nicht mehr vorstellbar ist. Die IT unterstützt sämtliche Aktivitäten, indem die geeigneten Anwendungen zur Verfügung gestellt werden, um einen reibungslosen Geschäftsbetrieb sicherzustellen. Dazu müssen mehr als 329.000 Mitarbeiter mit Informationen und Daten versorgt und über 30 Rechenzentren weltweit betrieben werden.

Damit die IT ihren Wertbeitrag auch nachhaltig erbringen kann, besteht die Zielsetzung der IT darin, die Voraussetzungen für eine erfolgreiche Umsetzung der Unternehmensstrategie zu schaffen [Mue06]. Eine Maßnahme, um diese Voraussetzungen zu schaffen, war und ist die Entwicklung und Umsetzung einer BI-Strategie.

Die BI-Strategie beschreibt im Sinne eines langfristigen und unternehmensweiten Master-Plans, wie das Unternehmen die Nutzung von Informationen ausgestaltet. Hierdurch schafft sie die Grundlagen für den Aufbau konzernweiter BI-Initiativen und Projekte.

Oberste Ziele der BI-Strategie sind die Optimierung der Gesamtkosten für die BI-Landschaft und die Ausweitung des Business Use für den Konzern (Abbildung 1). Voraussetzung für eine Kostenoptimierung ist zunächst die Schaffung von Transparenz in Bezug auf alle BI-Initiativen im Konzern. Auf Grundlage der geforderten Transparenz können konkrete Maßnahmen zur Schaffung einer standardisierten und harmonisierten Prozess- und Applikationslandschaft definiert werden. Nur eine solche Landschaft führt zu einer Senkung der Supportkosten und ermöglicht im Gegenzug größere Spielräume beim Umsetzen von Projekten, die den Wertbeitrag für das Unternehmen erhöhen. Eine Verbesserung des Business Use wird außerdem durch eine erhöhte Prozessunterstützung und eine höhere Prozessabdeckung sowie der Möglichkeit, Enterprise Reporting durchzuführen, erreicht [VW07].

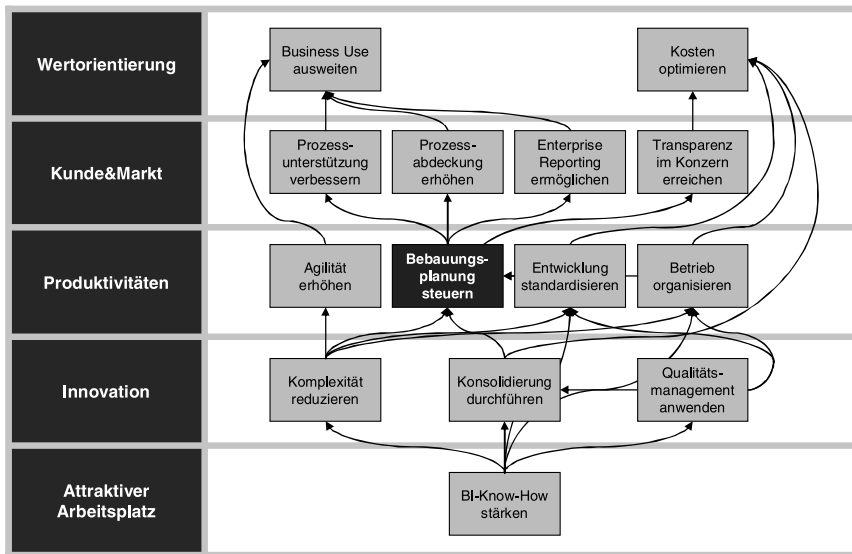


Abbildung 1: Strategielandkarte entlang der Dimensionen der Volkswagen Balanced Scorecard

Ein wichtiges Instrument zur Erreichung der genannten Ziele ist die gesteuerte BI-Bebauungsplanung im Rahmen der BI-Governance. Durch sie entsteht u. a. die benötigte Transparenz über den jeweils aktuellen Ist-Zustand der BI-Landschaft. Dieser hilft wiederum, die Aussagefähigkeit und die Qualität der BI-Bebauungsplanung zu verbessern. Die Konsolidierung der existierenden BI-Landschaft, die Standardisierung und Organisation von Entwicklung und Betrieb sowie eine reduzierte Komplexität der BI-Landschaft und BI-Anwendungen sind ebenfalls Bestrebungen, die sich in der BI-Bebauungsplanung niederschlagen. Durch eine reduzierte Komplexität wird die Agilität erhöht, was wiederum zusätzliche Möglichkeiten eröffnet, den Nutzen für die Fachanwender zu steigern [So08].

Gegenstand des nachfolgenden Kapitels ist zunächst die Beschreibung der wesentlichen Inhalte der BI-Strategie. Im Anschluss wird das BI-Bebauungsmanagement als Werkzeug der BI-Governance und als ein Mittel zur Umsetzung der BI-Strategie detailliert dargestellt.

2 Die BI-Strategie des Volkswagen-Konzerns

2.1 Motivation

Die Entwicklung und Umsetzung der BI-Strategie wurde aktiv aus der IT heraus getrieben. Dazu wurde bereits 2005 mit dem IT-Kompetenzfeld Business Intelligence (IT-KF BI) eine organisatorische Einheit geschaffen, um die die notwendigen fachlichen, organisatorischen und technischen Aktivitäten im Rahmen des BI-Programms zu koordinieren [So08].

Durch IT-Enabler wie BI wird es möglich, nachhaltiger Mehrwert für das Unternehmen durch systematische, bereichsübergreifende Zusammenführung und Nutzung von Informationen zu schaffen [Wi08]. Der Volkswagen-Konzern wird zukünftig hohe Anforderungen an die Verwendung und Analyse von Informationen zur Entscheidungsunterstützung auf strategischer, taktischer und operativer Ebene haben. Die nachstehend genannten Faktoren motivierten den Volkswagen-Konzern zur Entwicklung einer BI-Strategie für den Gesamtkonzern:

- Ein erster System-Überblick der Marken Volkswagen und Audi ergab ein heterogenes und komplexes Bild der BI-Landschaft, welches einem Flickenteppich ähnelte. Die BI-Landschaft war gekennzeichnet durch den Einsatz unterschiedlichster Produkte, Standards, Technologien und Architekturen. Eine Reduktion der Komplexität und die Hebung von Synergiepotentialen durch Konsolidierungsmaßnahmen speziell in den Ebenen der Infrastruktur, der Daten und Datenhaltungen sowie der Applikationen sind daher erforderlich.
- Der Volkswagen-Konzern operiert in einem durch hohe Dynamik und Wettbewerbsintensität geprägten internationalen und komplexen Umfeld, in dem Entscheidungen durch verlässliche Informationsgrundlagen unterstützt werden müssen. Dazu müssen alle BI-Systeme unabhängig von ihrer Reichweite und ihrem Einsatzspektrum in ihrem Gültigkeitsbereich optimal mit Daten versorgt werden. Ziel ist, entsprechend den Grundprinzipien der Informationslogistik, die richtige Information in der richtigen Qualität und Sicherheit zur richtigen Zeit für den richtigen Adressaten am richtigen Ort zur Verfügung stellen zu können.
- Der Volkswagen-Konzern hat sich in den vergangenen Jahren zu einem prozess-orientierten Unternehmen entwickelt. So kann den Anforderungen der dynamischen Märkte und deren regulatorischen Vorgaben zeitnah begegnet werden. Zahlreiche Entscheidungen werden zukünftig dezentral getroffen. Für die Gesamtsteuerung des Unternehmens ist es daher unerlässlich, zentrale integrierte Kennzahlensysteme aufzubauen und die informationstechnischen Versorgungsströme zu etablieren. [So08].
- Die IT des Volkswagen-Konzerns bewegt sich heute in einem Spannungsfeld. Auf der einen Seite steht eine zunehmende Komplexität der Aufgaben, die in neuen Anforderungen an die IT-Lösungen zum Ausdruck kommt. Auf der anderen Seite ist ein signifikanter Kostendruck zu spüren, der direkt auf IT-Budgets weitergegeben wird. Der Erklärungsdruck bezüglich der Notwendigkeit von Investitionen lastet dabei nicht nur auf den beauftragenden Fachbereichen, sondern auch auf der IT [Mue06]. Investitionen in IT-Systeme werden sorgfältig geprüft und einer Wirksamkeitsbetrachtung unterzogen. Eine übergeordnete strategische Gesamtplanung für BI ist daher erforderlich, um die Investments in BI aus Unternehmenssicht effektiv und effizient zu gestalten.

2.2 Ziele

Die BI-Strategie des Volkswagen-Konzerns entstand in einem strukturierten Entwicklungsprozess. Dieser Strategieentwicklungsprozess ist zunächst BI-unspezifisch, die BI-Besonderheiten kommen jedoch in der konkreten Ausgestaltung der einzelnen Prozessschritte Analyse, Design, Umsetzung und Controlling zum Tragen [TZ06].

Abgeleitet wurde die konzernweit definierte BI-Strategie aus einer stringenten Zielhierarchie, innerhalb der die BI-Strategie in die IT-Strategie integriert ist. Letztere leitet sich wiederum aus der Konzernstrategie ab, so dass die auf den jeweils einzelnen Ebenen der Strategiehierarchie relevanten Kernaussagen auch für die BI-Strategie Gültigkeit haben und den Rahmen der BI-Strategie abstecken. So gewährleistet der strukturierte Strategieentwicklungsprozess, der zur Erstellung und Weiterentwicklung der BI-Strategie führt, das Business-IT-Alignment bezogen auf den Einsatz von Business Intelligence im Volkswagen-Konzern.

„BI – der Weg zum Wissensvorsprung des Volkswagen-Konzerns. Einsichten schaffen, Informationen integrieren, Kompetenzen vernetzen“ so lautet die Vision des Volkswagen-Konzerns für Business Intelligence. Diese Vision beinhaltet als einen wesentlichen Teil die Integration von Informationen des gesamten Volkswagen-Konzerns und dessen Umfelds über die Kernprozesse hinweg zu einer unternehmensweiten Informationsplattform. Darüber hinaus soll eine Vernetzung von Kompetenzen durch Querschnittsfunktionen sowohl prozess-, als auch marken- und regionsübergreifend vorangetrieben werden. Ein weiterer Teil der Vision ist es, tiefere Einsicht in Daten (Unternehmenszustände) zu schaffen, auf deren Basis Entscheidungsgrundlagen für operative Prozesse entstehen. Alle diese Teile zusammen bewirken den angestrebten Wissensvorsprung des Volkswagen-Konzerns. Abgeleitet ergibt sich die Mission der BI-Strategie in den folgenden vier Feldern:

- Erschließung neuer fachlicher Möglichkeiten: Das BI-Programm verstärkt den IT-Bereich als Innovator im Unternehmen.
- Effizienz durch Standardisierung: Das BI-Programm überführt die BI-Landschaft in die BI-Plattform durch die Definition von BI-Standards und deren Implementierung in den Bereichen BI-Governance, Plan, Build und Run.
- Qualität vor Quantität: Das BI-Programm steht im Unternehmen für die fachliche, organisatorische und technische Qualität der BI-Landschaft.
- Mehr BI fürs Geld: Das BI-Programm schafft die Voraussetzungen, um den steigenden Anforderungen an BI mit den zur Verfügung stehenden Mitteln zu begegnen.

Aus Vision und Mission wurden die zu erreichenden vier Kernziele „Kosten pro BI-System senken“, „Qualität von BI erhöhen“, „Standardisierung erhöhen“ und „Innovativität von BI erhöhen“ abgeleitet [So08].

2.3 Umsetzung

Im Rahmen einer übergreifenden BI-Strategie sind Verknüpfungen zu Detailstrategien definiert:

- In der fachlichen BI-Strategie werden BI-Konzepte für die Fachbereiche erarbeitet.
- In der BI-Produktstrategie werden Kriterien für strategische Lösungspartnerschaften definiert, um Grundlagen zu schaffen, die Handlungsfelder von BI mit Standardprodukten zu hinterlegen.
- Die BI-Konsolidierungsstrategie beschreibt die mit dem Konsolidierungsvorhaben verbundenen strategischen Vorgaben (Planung, Steuerung, Durchführung).
- Die notwendigen Voraussetzungen für eine optimale Wertschöpfung durch Business Intelligence werden in der BI-Daten-Managementstrategie geschaffen.
- Vorgaben für die Bereitstellung einer standardisierten, zukunftssicheren Infrastruktur für die BI-Plattform sind in der BI-IT-Services-Strategie enthalten.
- Die BI-AMS-Strategie (Application Management Services) beschreibt den Aufbau eines BI-AMS-Clusters zum Betrieb dieser BI-Plattform [So08].

Das BI-Programm wurde als konzernübergreifende Initiative, mit dem Auftrag die BI-Aktivitäten im Konzern zu steuern und zu bündeln, eingeführt und ist als konzernübergreifender Rahmen für die notwendigen Maßnahmen zur Umsetzung der BI-Strategie auf allen Ebenen etabliert.

Das BI-Programm bildet die gemeinschaftliche Klammer über IT und Fachbereiche. Dabei hängt der Erfolg des BI-Programms maßgeblich von der aktiven Zusammenarbeit der beteiligten Bereiche ab. So sind die IT-Bereiche für die kostenoptimale Ausgestaltung, die Neu-/Weiterentwicklung von BI-Systemen sowie deren Betrieb verantwortlich. Die Fachbereiche hingegen verantworten die Definition der fachlichen Anforderungen und das Heben von Nutzenpotentialen. Die Zusammenarbeit im BI-Programm erfolgt im Rahmen von Projekten und wird durch eine geeignete Gremienlandschaft koordiniert [So08].

Treiber des BI-Programms ist das IT-Kompetenzfeld Business Intelligence, das einen organisatorischen Rahmen vorgibt. Das IT-Kompetenzfeld Business Intelligence zentralisiert wesentliche Aufgaben über die Funktionen des BI-Competence Centers, des zentralen BI-Datenmanagements und der BI-Governance. Letztere ist Gegenstand der weiteren Betrachtung.

3 Die BI-Bebauungsplanung als Instrument der BI-Governance

3.1 BI-Governance

Ein bestehendes IT-Governance-Modell sollte, sofern dieses im Unternehmen vorhanden ist, die zentrale Referenz für die Ausgestaltung einer BI-Governance sein. Die BI-Governance ergänzt und konkretisiert die IT-Governance dabei lediglich um BI-spezifische Ableitungen. Bei der konkreten Ausgestaltung der BI-Governance im Volkswagen-Konzern konnte nicht auf ein bestehendes IT-Governance-Modell innerhalb des Konzerns zurückgegriffen werden. Daher wurde unter Berücksichtigung der in der einschlägigen Literatur vorzufindenden, aber uneinheitlichen Definitionen und Modelle [GJFG07] eine eigene Sicht auf die BI-Governance definiert.

So übt die BI-Governance im Volkswagen-Konzern eine Servicefunktion aus und unterstützt die zentralen Umsetzungsschritte zur sukzessiven Einführung der BI-Strategie. Darüber hinaus verantwortet die BI-Governance die Vollständigkeit, Passgenauigkeit und Redundanzfreiheit aller damit verbundenen und übergeordneten Maßnahmen.

Die BI-Governance umfasst dazu die im Unternehmen notwendigen Steuerungsfähigkeiten zur Formulierung, Umsetzung und Kontrolle der BI-Strategie. Hierzu gehören insbesondere die Sicherstellung transparenter Entscheidungsgrundlagen für das Management, die Methoden zur Steuerung der bestehenden BI-Landschaft, die Ausrichtung der Governance-Instrumente auf die Überführung der BI-Landschaft in die BI-Plattform und deren strategiekonformen Aufbau und Betrieb. Kernprozessübergreifende Anforderungen an BI, bestehende oder geplante BI-Systeme, die in mehreren Prozessen zum Einsatz kommen und Entscheidungen über Prozessgrenzen hinweg unterstützen, können ebenso wie die durchgängige Gestaltung der gesamten BI-Plattform durch die BI-Governance koordiniert umgesetzt werden. Sie trägt damit zur Sicherstellung der Fusion von BI und Geschäft bei, in dem Maße, wie BI darauf wirken kann.

Die Ziele der BI-Governance sind:

- Etablierung von BI-Governance-Prozessen (und -Objekten) und einer abgestimmten Organisation sowie klar definierter und transparenter Entscheidungsrechte und -pflichten. Damit wird eine ganzheitliche Gestaltung und Steuerung der BI-Landschaft gemäß der BI-Strategie – konzernweit und geschäftsprozessübergreifend – ermöglicht.
- Sicherstellung der Umsetzung der vorgegebenen Standards in den Projekten (BI-Lifecycle-Betrachtung, Anwendung des Volkswagen Systementwicklungsprozesses (SEP), Bausteinverwendung, Datenmanagement, Betrieb von BI-Systemen).
- Definition eines abgestimmten Sets von Instrumenten (Governance-Objekten) und Prozessen [So08].

Als ein wesentliches Instrument steht die BI-Bebauungsplanung im Zentrum der BI-Governance. Diese stellt zusammen mit dem Portfolio-Instrumentarium Werkzeuge zur Überführung des aktuellen Ist-Zustands (der „BI-Landschaft“) in den wohldefinierten Soll-Zustand (der „BI-Plattform“) im Sinne der BI-Strategie bereit.

Die BI-Landschaft umfasst dabei alle IT-Systeme, die BI-Technologien nutzen oder beinhalten. Die BI-Plattform als Soll-Vorstellung besteht aus ausgewählten Data Warehouses (den sogenannten Gravitationszentren) und BI-Services, die den Vorgaben der BI-Strategie entsprechen. Die Gravitationszentren sind als konsolidierte Datenbasis die tragenden Säulen des BI-Informationsangebotes des Volkswagen-Konzerns. Die BI-Services haben die Aufgabe, die fachlichen Anforderungen zu unterstützen, in dem die Daten der Gravitationszentren anwendungs- bzw. auswertungsbezogen aufbereitet und in einen betriebswirtschaftlichen Kontext gestellt werden. In den Geschäftsprozessen werden lediglich BI-Services im Sinne dispositiver Applikationen sichtbar, die Gravitationszentren als zu Grunde liegende Datenbasis sind vollständig gekapselt. Durch diese Entkopplung ist es möglich, den Zielkonflikt zwischen Flexibilität und Agilität seitens der BI-Services einerseits und Qualität und Standardisierung auf Ebene der Gravitationszentren andererseits aufzulösen. Die BI-Plattform ist als verbindlicher gestalterischer Rahmen im Volkswagen-Konzern nur über einen koordinierten Integrationsprozess umsetzbar [So08]. Dieser Integrationsprozess ist dafür verantwortlich, dass die Vorgaben der BI-Strategie eingehalten werden. So hat die Definition der BI-Services-Schicht und der Gravitationszentren-Schicht zur Konsequenz, dass BI-Systeme in diese beiden Anteile zerlegt werden, um die Anteile anschließend der jeweiligen Schicht zuzuführen. Für neue BI-Systeme erfolgt dies grundsätzlich während der initialen Erstellung. Vorhandene Systeme werden im Rahmen eines allgemeinen Konsolidierungsprozesses oder während der normalen Weiterentwicklung business-case-getrieben angepasst.

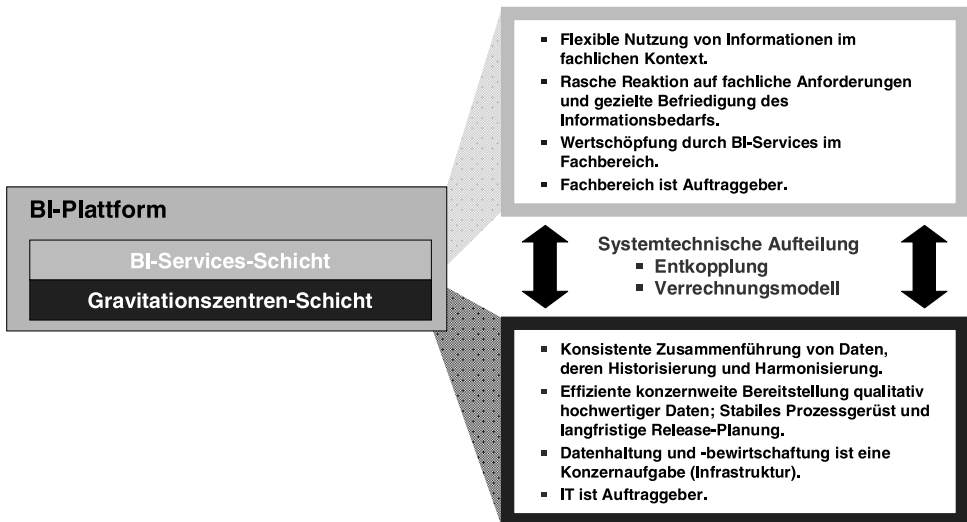


Abbildung 2: Aufgaben und Zielsetzungen der BI-Plattform

3.2 Motivation für eine BI-Bebauungsplanung

Die Motivation für die Einführung einer BI-Bebauungsplanung fußt auf einer BI-Landschaft, die durch eine hohe Komplexität und Heterogenität gekennzeichnet ist. Die Beherrschbarkeit der BI-Landschaft ist im Volkswagen-Konzern zum einen durch die hohe Anzahl von BI-Systemen (derzeit sind über 200 BI-Systeme bekannt) und zum anderen durch die Vielzahl der eingesetzten Architekturen, Prozesse und Technologien stark beeinträchtigt. Die Folgen dieser Beeinträchtigungen sind erhöhter Aufwand bei der konzernweiten Planung und Umsetzung eines gezielten und effizienten Einsatzes von BI sowie die Gefahr von Redundanzen, Überschneidungen und schließlich einer mangelnden Geschäftsunterstützung durch BI.

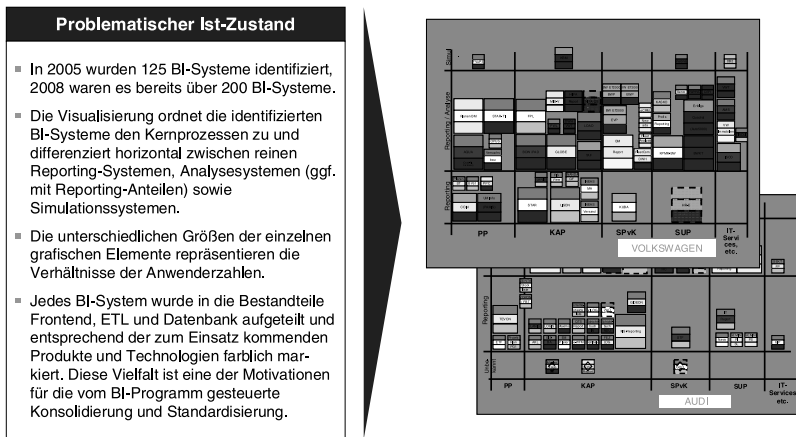


Abbildung 3: BI-Landschaft

Die Etablierung einer BI-Bebauungsplanung ist demnach kein Selbstzweck, sondern stellt mit dem BI-Bebauungsplan ein Werkzeug dar, das langfristig zur Umsetzung der Unternehmensziele und der in der BI-Strategie definierten Ziele bezüglich Harmonisierung und Standardisierung, Reduzierung der Systemvielfalt und des Schließens von Prozesslücken dient. Der BI-Bebauungsplan schafft die notwendige Transparenz und bildet dadurch die Grundlage für Entscheidungsfindung und Steuerung der BI-Governance im Rahmen der Ausgestaltung der BI-Plattform. Als Diskussions- und Planungsinstrument im Konzern gibt der BI-Bebauungsplan Orientierung für die künftige Gestaltung der BI-Plattform zur optimalen Unterstützung der Fachbereiche und der sich permanent wandelnden Prozesse.

Als Planungsunterstützung steht der BI-Bebauungsplan begleitend zu den jährlichen Planungsrunden zur Verfügung. Er gewährleistet damit die Berücksichtigung der mittel- und langfristigen Ziele des Konzerns bei der Verteilung des jährlichen IT-Budgets. Die regelmäßige Überarbeitung des Bebauungsplans gewährleistet, dass neue Anforderungen sowohl hinsichtlich Strategie-, Prozessveränderungen und wirtschaftlicher Vorgaben als auch aufgrund neuer Systeme und Technologien berücksichtigt werden. Dieses führt dazu, dass der Umfang, die Konkretisierung und die Qualität des BI-Bebauungsplans bei den jährlichen Aktualisierungen zunehmen.

3.3 Ziele der BI-Bebauungsplanung

Mit Bezug auf die Motivation für die Etablierung einer BI-Bebauungsplanung wurden folgende Ziele der strategischen Maßnahme BI-Bebauungsplanung durch die BI-Governance definiert.

- Schaffung von Transparenz über die bestehende BI-Landschaft.
- Erfassung und Visualisierung des Ist-Zustandes der BI-Landschaft unter Verwendung unterschiedlicher, zweckorientierter Sichten.
- Erarbeitung einer Soll-Bebauung, abgeleitet aus den übergeordneten Strategien und in Zusammenarbeit mit allen Stakeholdern.
- Ableitung, Auswahl, Priorisierung, Kombination und Steuerung der notwendigen Maßnahmen aus der Überführung der BI-Landschaft in die BI-Plattform.
- Identifizierung möglicher und neuer Projekte aus diesen Maßnahmen und Einbringung in die jährlichen Planungsrunden.
- Integration neuer fachlicher Anforderungen in die Gestaltungs- und Planungsprozesse.
- Ganzheitliche, prozessübergreifende Gestaltung der BI-Landschaft mit dem Ziel der Überführung in die BI-Plattform auf den Ebenen der BI-Services und Gravitationszentren.
- Klärung der wechselseitigen Abhängigkeiten zu den im Konzern existierenden Bebauungsplänen, Portfolios und Geschäftsprozessen.

3.4 Umsetzung der BI-Bebauungsplanung

Konzept und Vorgehen

Die BI-Bebauungsplanung orientiert sich an dem Konzept der Unternehmensarchitektur. Unternehmensarchitekturen (engl. Enterprise Architecture) bieten eine integrierende Sicht auf Unternehmen, welche insbesondere organisatorische und informationstechnische Aspekte in Zusammenhang zueinander setzt. Diese integrierende Sicht ermöglicht u. a. Architekturtransformationen zu planen und zu managen. Aus der Betrachtung unterschiedlicher Ansätze zur Beschreibung von Unternehmensarchitekturen wie von Krömar, Foegen, Dern oder Hafner/Schelp/Winter wird deutlich, dass alle Architekturansätze eine ganzheitliche Sichtweise, ausgehend von der Strategie über organisatorische und fachliche Aspekte zu technischen Fragestellungen, haben [SO07]. Diese ganzheitliche Sichtweise auf die BI-Landschaft bildet die Grundlage für die BI-Bebauungsplanung. Das Konzept und das Vorgehen der BI-Bebauungsplanung verdeutlicht die folgende Grafik:

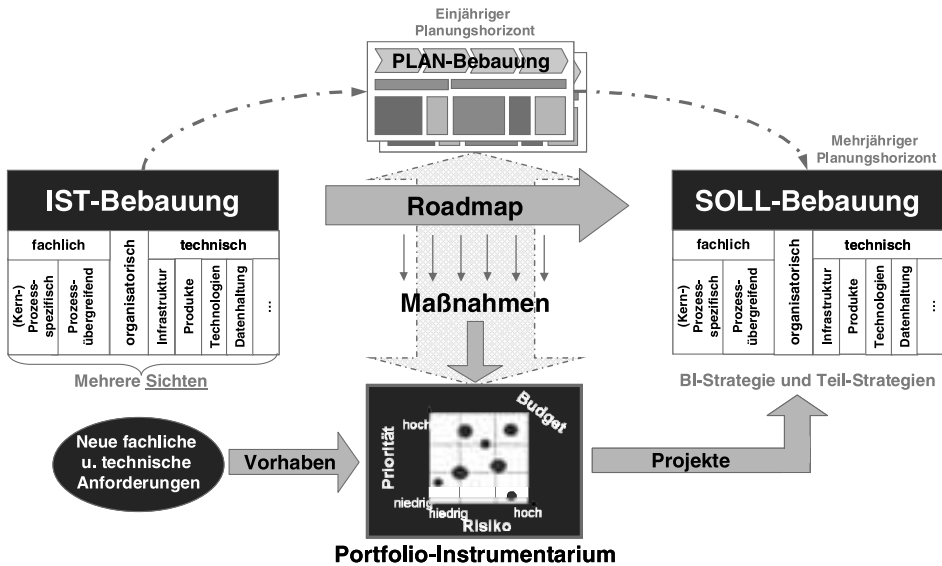


Abbildung 4: Konzept der BI-Bebauungsplanung

Nach Klärung der generellen Rahmenbedingungen für das BI-Bebauungsmanagement (3.4.2) wurden die vorhandenen BI-Systeme inventarisiert und systematisch anhand fachlicher, technischer und organisatorischer Merkmale strukturiert (3.4.3). Die erhobenen Informationen wurden in einem zentralen Repository abgelegt, analysiert und zweckorientiert visualisiert. Für die Darstellung der BI-Bebauung – unabhängig davon, ob als Ist-Zustand oder als Soll-Vorstellung – ist die sinnvolle und konsistente Zusammenstellung und Erstellung mehrerer unterschiedlicher Visualisierungsformen erforderlich. Das ist notwendig, da unterschiedliche Aspekte fachlicher, organisatorischer und technischer Perspektiven eines BI-Systems nicht in einer einzigen Darstellung abgebildet werden können, sondern mehrere Sichten benötigt werden. Darüber hinaus besteht die Notwendigkeit, für die Ergebnisse der Bebauungsplanung eine Form der Darstellung zu finden, die so weit wie möglich verdichtet und gleichzeitig noch die notwendige Detaillierung der Inhalte erlaubt. Zur Visualisierung werden im Rahmen der Bebauungsplanung Softwarekarten als graphische Modelle der BI-Landschaft eingesetzt [LMW05]. Die Ergebnisse werden im BI-Bebauungsplan dokumentiert.

Aus der BI-Strategie und deren Teilstrategien wird die Soll-Bebauung (BI-Plattform) in Zusammenarbeit mit allen Stakeholdern der BI-Landschaft entwickelt. Der mehrjährige Planungshorizont erfordert eine rollierende Anpassung, Ergänzung und Verfeinerung der Soll-Bebauung (3.4.4). Der Transformationsprozess vom jeweiligen Ist zum Soll wird durch eine Roadmap geeigneter Maßnahmen repräsentiert. Diese Maßnahmen fließen gemeinsam mit den aus neuen fachlichen Anforderungen resultierenden Vorhaben in eine Portfolio-Betrachtung ein. Diese Betrachtung dient dazu, die durchzuführenden Projekte zu identifizieren, zu priorisieren und zu kombinieren. Anhand der Projekte eines Planungsjahres lässt sich dann die Plan-Bebauung ableiten, die am Ende des Planungsjahres erreicht werden soll. Die identifizierten Projekte finden Eingang in die Budgetplanungsrunden des jeweiligen Jahres (3.4.5).

Da sich die BI-Landschaft permanent durch äußere und innere Einflüsse weiterentwickelt, sind eine kontinuierliche Fortschreibung, Anpassung und Weiterentwicklung der BI-Bebauungsplanung zwingend erforderlich.

Rahmenbedingungen

Unter der Prämisse „*Think big, start small*“ wird die operative Umsetzung des BI-Bebaungsmanagements zunächst von einem kleinen Team verantwortet. Erfolgsfaktoren dieses Teams sind umfassendes Wissen über den Volkswagen-Konzern und dessen Prozesse sowie Kenntnisse von Methoden zu Aufbau, Gestaltung und Visualisierung von Bebauungsplänen.

Zur konkreten Einführung des BI-Bebaungsmanagements wurden zunächst die Rahmenbedingungen festgelegt. Dazu zählten neben der Identifikation der Stakeholder und der finanziellen Absicherung des Vorhabens auch die Definition der notwendigen Prozesse sowie die Festlegung der Aufgaben und Rollenverteilungen innerhalb des BI-Bebaungsmanagements. Darüber hinaus wurde die Ausgestaltung eines planmäßigen Vorgehens zur Erstellung des BI-Bebauungsplanes erarbeitet und die Zusammenarbeit bzw. Abgrenzung des BI-Bebaungsmanagements vom Konzern-Bebaungsmanagement geregelt.

Das Konzern-Bebaungsmanagement ist mitverantwortlich für die Erstellung der Bebauungspläne – den sogenannten Master Construction Plans (MCPs) – für die vier Kerngeschäftsprozesse des Volkswagen-Konzerns. Bei den Kerngeschäftsprozessen handelt es sich im Detail um:

- den Produktprozess (PP), der den gesamten Lebenszyklus eines Produkts, vom Beginn der Produktplanung und -definition über die Produktentstehungsphase bis zur Serienbetreuungsphase, umfasst.
- den Kundenauftragsprozess (KAP), der die Fertigung eines Fahrzeugs im Kundenauftrag, die gesamte Auftragsverfolgung vom Produktprozess über den Planungsprozess, die Auftragseinplanung, die Fertigungssteuerung sowie die Distribution bis zur Fahrzeugübergabe beinhaltet.
- den Serviceprozess vor Kunde (SPK) als Prozess, der das gesamte Kundenbeziehungs-Management der Bereiche Originalteile, Werkstattservice, Gebrauchtwagen sowie Finanzdienstleistungen inklusive Leasing und Vermietung umfasst.
- Steuerungsprozesse und unterstützenden Prozesse (SUP) mit den Konzernbereichen wie Unternehmenssteuerung, Personal, Finanzwesen/Rechnungswesen/Controlling, Qualitäts-Management und allgemeine Dienstleistungen.

Der für jeden Kernprozess erstellte Master Construction Plan verdeutlicht, welche IT-Anwendungen auf Grundlage der Konzernstrategie und relevanter Kernprozess- bzw. Geschäftsbereichsstrategien die Anforderungen der einzelnen Kernprozesse erfüllen bzw. zukünftig erfüllen werden. So wird u. a. sichtbar, welche Konzernsysteme kurz- oder mittelfristig abgelöst werden, durch welche Nachfolgesysteme (sofern bekannt)

diese Systeme ersetzt werden oder welche Geschäftsprozesse zukünftig durch welche Systeme unterstützt werden. Die MCPs sind somit auch ein Kommunikationsmittel zwischen IT- und Fachbereich in den jeweiligen Kernprozessen. Eine Harmonisierung der MCPs über die Kernprozessbereiche hinweg steht noch aus.

Der BI-Bebauungsplan ergänzt die für die Kerngeschäftsprozesse existierenden Master Construction Plans um die erforderliche prozessübergreifende Steuerungssicht.

Damit das BI-Programm seinen Gestaltungsauftrag, die BI-Landschaft in die BI-Plattform zu überführen, wahrnehmen kann, ist eine kernprozessübergreifende Sicht unerlässlich. Die bestehenden MCPs legen ihren Fokus auf die Geschäftsprozesse und die Identifikation der Systeme, die diese Prozesse unterstützen. Das BI-Programm hat eine andere Sicht auf operative Systeme, da diese nicht Objekt der Planung und Gestaltung sind, wohl aber in ihrer Funktion als Datenlieferanten Relevanz haben. Im Fokus des BI-Programms stehen die dispositiven Systeme und die prozessübergreifenden Informationsflüsse. In der BI-Perspektive geht es u. a. um die Fragen, welche Daten in den einzelnen Geschäftsprozessen anfallen, in welchen BI-Systemen und für welche Anwender diese Daten vorgehalten werden oder welche Informationen zur Prozess-(übergreifenden)Steuerung notwendig sind.

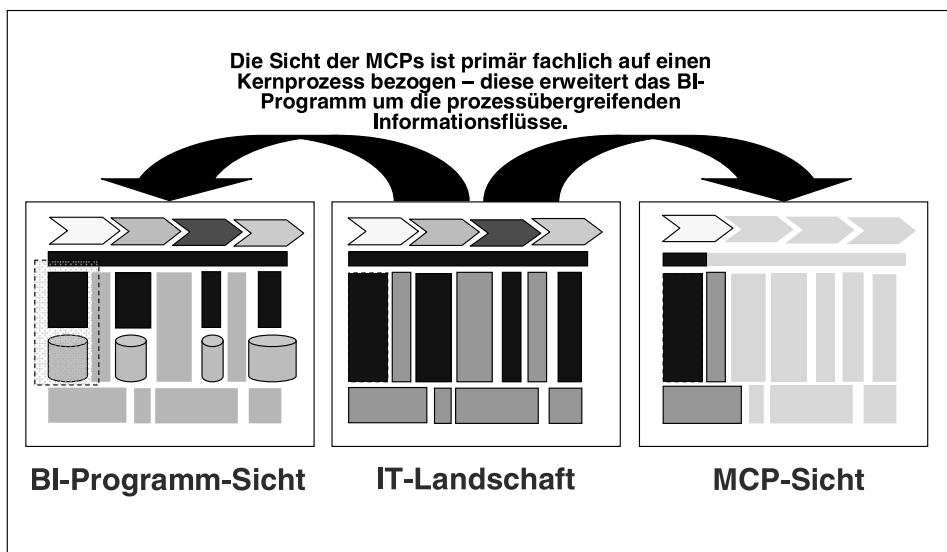


Abbildung 5: MCP- und BI-Programm-Sicht auf die IT-Landschaft

Dementsprechend sind Daten und Datenströme wesentliche Elemente des BI-Bebauungsplans.

Die an der Gestaltung der IT-Landschaft im Volkswagen-Konzern beteiligten Instanzen haben unterschiedliche, sich jedoch ergänzende Sichten auf das Gesamtgebilde aus Geschäftsprozessen und die dazugehörige Informationstechnologie.

Informationserhebung

Nach Klärung der Rahmenbedingungen für die Erstellung des BI-Bebauungsplanes wurde festgelegt, wie detailliert der BI-Bebauungsplan auszugestalten ist und welche Informationen dazu erforderlich sind. So orientiert sich die BI-Bebauung auf Prozessebene analog zu den MCPs primär an den Prozessen der Ebene 3. Auch wird die Sicht auf Funktionsmodule, als quer zu den Prozessen liegende funktionale Anforderungen, verwendet. Diese Module ordnen in den MCPs und der BI-Bebauung die Anwendungen in Gruppen und gewährleisten gleichzeitig eine sinnvolle Darstellungsform für die anstehende Diskussion und Abstimmung im Konzern. Zusätzlich kann die Zuordnung der operativen Systeme zu den Prozessen den MCPs entnommen werden. Um jedoch der beschriebenen BI-Sicht gerecht zu werden, wurde als erster Schritt im Rahmen einer Ist-Erhebung ein Überblick geschaffen:

- welche BI-Systeme und -Projekte im Konzern vorhanden sind
- welche Funktionen und Fachlichkeit sie abdecken
- welchem Kerngeschäftsprozess sie zuzuordnen sind
- welche Geschäftseinheit bzw. welche Person für sie verantwortlich ist
- welche Schnittstellen existieren
- welche Technologien eingesetzt werden.

Diese Informationen werden im BI-Portfolio, welches z. Zt. ca. 200 BI-Elemente umfasst, gesammelt und aktuell gehalten. Als BI-Elemente werden im Volkswagen-Konzern alle existierenden (betriebs-) IT-Systeme, die zur BI-Landschaft gezählt werden, sowie Projekte, die die Erstellung, Weiterentwicklung oder Abschaltung von entsprechenden IT-Systemen der BI-Landschaft zum Ziel haben, bezeichnet. Erhoben wurden die Informationen durch eine Analyse des Konzernsystemverzeichnisses, mit Hilfe von Experten (u. a. den Konzern-Architekten) und dem Auslesen von Metadaten der eingesetzten BI-Werkzeuge.

Im zweiten Schritt wurden in Interviews mit den jeweiligen System- und Prozessverantwortlichen auf Sachbearbeiter-Ebene bzw. im mittleren Management die in den operativen Systemen vorhandenen Daten und deren Zuordnungen zu den entsprechenden Prozessen, BI-Systemen und Datenverwendern betrachtet.

Um die Handhabbarkeit der Informationen über Daten und Datenströme zu gewährleisten, wurden sie nicht auf Attributebene betrachtet, sondern als fachlich identifizierte Oberbegriffe (z. B. *Fahrzeug*, *Lieferschein*, *Werk*) in Form von Datenklassen zusammenfassend konkretisiert. Diese generisch definierten Datenklassen werden später durch das BI-Datenmanagement weiter verfeinert¹.

¹ Diese Schritte werden hier jedoch nicht weiter betrachtet.

In den Interviews wurden außerdem die zukünftigen fachlichen, technischen und organisatorischen Herausforderungen der befragten Bereiche erarbeitet. Deutlich wurden dabei u. a. die fachbereichsübergreifend starke Nachfrage nach Finanz- und Qualitätsdaten, die starken Redundanzen bei der Speicherung von zentralen Datenklassen oder Performance-Optimierungspotentiale.

Alle diese Grundlagen werden im weiteren Verlauf zur Analyse der BI-Landschaft und zur Ausgestaltung des BI-Bebauungsplans herangezogen.

Erstellung der Ist und Soll-Bebauung

Im Rahmen der Visualisierung der Ist- und Soll-Bebauung steht nicht die komplette Visualisierung der BI-Landschaft, sondern die Darstellung wesentlicher Informationen für ein konkretes Untersuchungsfeld im Vordergrund. Die Visualisierung hilft vorrangig bei der Verdeutlichung konkreter Sachverhalte, bei Diskussionen mit den Stakeholdern der BI-Bebauung im Allgemeinen und insbesondere bei der Entscheidung hinsichtlich der einzuleitenden nächsten Schritte.

Zur Visualisierung dienen in der aktuellen Version des BI-Bebauungsplans primär die Prozess-System-Datenlandkarte und die Schnittstellendiagramme. Letztere setzen die wesentlichen BI-Systeme zueinander in Beziehung.

Materialbeschaffung

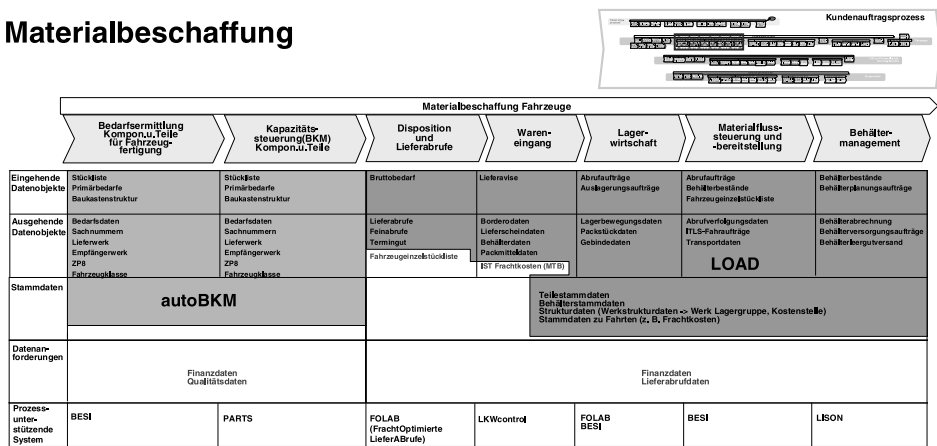


Abbildung 6: Beispiel Prozess-System-Datenlandkarte

In der Prozess-System-Datenlandkarte werden die Systeme anhand des von ihnen unterstützten Prozesses und Funktionsmoduls verortet sowie die in diesen Systemen und Prozessabschnitten eingehenden, anfallenden und ausgehenden Datenklassen visualisiert. So kann verdeutlicht werden, welche Daten und Informationen in den jeweiligen Prozessen und Systemen entstehen und benötigt werden. Darüber hinaus wird dargestellt, in welchen BI-Systemen diese Daten abgelegt sind und welche Daten zurzeit noch nicht für dispositive Auswertungen zur Verfügung stehen. Ergänzt wird die Darstellung durch Daten und Informationen, die für eine verbesserte Fachbereichsunterstützung hilfreich

wären, dem Fachbereich aber zurzeit noch nicht zugänglich sind bzw. noch nicht zur Verfügung stehen.

Die Prozess-System-Datenlandkarte hilft in ihrer Matrixdarstellung aus Systemen, Modulen, Daten, Prozessen und Datenwünschen, die anstehende Diskussion und Abstimmung im Konzern zielgerichtet voranzutreiben. Die gewählte Darstellungsform macht noch auszubauende Bereiche sichtbar, zeigt Lücken in der Prozessunterstützung ebenso wie Überschneidungen in der Prozessabdeckung oder in den Datenhaltungen

Darüber hinaus stellt die Prozess-System-Datenlandkarte die Datenverwender (aktuell und zukünftig) gemeinsam mit den Datenerzeugern dar. Mit Hilfe dieser Darstellung können alle Beteiligten (Datenverwender, Datenerzeuger, Datenmanagement) gesteuert von der BI-Governance eine zielgerichtete Datendistribution gemäß der BI-Informationslogistik ausgestalten.

Für die Erstellung der Soll-Bebauung wird ähnlich wie für die Erhebung und die Darstellung der Ist-Bebauung ein evolutionäres Vorgehen gewählt. Die Darstellung der Soll-Bebauung erfolgt für den aktuellen BI-Bebauungsplan daher zunächst im Rahmen von bereits absehbaren Projekten, auf Grundlage der in den Gesprächen geäußerten Änderungswünsche hinsichtlich der BI-Landschaft und stets in enger Abstimmung mit den MCPs der Kernprozesse. Horizont der Soll-Bebauung ist zunächst ein Zeitraum von ein bis drei Jahren.

Bei der Erstellung der Soll-Bebauung spielt vor allem die Unterscheidung von Gravitationszentren und BI-Services eine Rolle. Primäres Ziel für die Soll-Bebauung im Bereich der Gravitationszentren ist die größtmögliche Abdeckung der Datenanforderungen sowie das sinnvolle „Schneiden“ der Gravitationszentren nach fachlichen, organisatorischen und technischen Gesichtspunkten.

Für die Gravitationszentren ist darüber hinaus relevant, welche Anteile globale Bedeutung haben und welche lokal für einzelne Prozesse gültig sind. Die globalen Anteile werden in einem konzernweiten (Informations-) Nukleus zusammengefasst. Der Nukleus teilt die jeweiligen Gravitationszentren in Gültigkeitsbereiche auf.

Der unternehmensweit vereinheitlichte Informationsnukleus wird in den kommenden Jahren iterativ entstehen und trägt fachlichen Anforderungen wie z. B. einer ganzheitlichen Kundensicht ebenso Genüge wie der allgemeinen Anforderung, über BI-Services qualitativ hochwertige Informationen zur Steuerung des Unternehmens anzubieten.

Ergänzt wird der konzernweit gültige Nukleus durch lokal gültige Informationen, die den einzelnen Regionen, Märkten, Werken und Organisationseinheiten die notwendige Flexibilität und Agilität ermöglichen. Die Prozesse der BI-Governance steuern die unterschiedliche Handhabung globaler und lokaler Anteile der Gravitationszentrenschicht und koordinieren die Gesamtentwicklung über das BI-Projekt-Portfolio. Der Nukleus unterliegt einem strengen Governance-Prozess. Seine Struktur muss sehr stabil gehalten werden. Änderungen des Nukleus sind daher aufwendig und abstimmungsintensiv.

Die Abdeckung der fachlichen Anforderungen ist dagegen das Ziel für die Aufstellung einer Soll-Bebauung im Bereich der BI-Services. Die Anforderungen können aus den für die Steuerung der einzelnen Fachbereiche bzw. der fachbereichsübergreifenden Steuerung maßgeblichen Kennzahlen, Bereichsstrategien, Prozessen und Benutzergruppen abgeleitet werden.

Die Entwicklung eines kompletten Bildes der BI-Soll-Bebauung für den Volkswagen-Konzern ist die zentrale Herausforderung der BI-Governance der kommenden Jahre.

Ableitung von Handlungsbedarfen und Maßnahmen

Nach Erstellung der Ist- und Soll-Bebauung folgt eine zielgerichtete Analyse, um Schwachstellen in der BI-Landschaft zu identifizieren und Verbesserungspotentiale aufzuzeigen. Dabei werden neben den Abhängigkeiten zwischen den Systemen, den Schnittstellen, der Konformität zur Referenzarchitektur des einzelnen BI-Systems auch die Prozessunterstützung oder Redundanzen und Lücken in der BI-Bebauung untersucht. Aus den Ergebnissen der Analyse lassen sich wiederum Handlungsempfehlungen bzw. Maßnahmen zur Überführung der BI-Landschaft in die BI-Plattform ableiten. Diese Maßnahmen bilden gemeinsam mit den aus den Interviews und den IT-Strategien resultierenden fachlichen Anforderungen und den daraus abgeleiteten Vorhaben einen Katalog von Handlungsempfehlungen.

Fokus der Tätigkeiten des BI-Bebauungsmanagements im ersten Jahr ist die Ableitung von konkreten Handlungsempfehlungen, die in den kommenden Jahren zur Umsetzung kommen sollen. Die Handlungsempfehlungen dokumentieren,

- welche Veränderungen an der bestehenden BI-Landschaft vorgenommen werden sollten, um die Geschäftsziele besser als bisher zu unterstützen
- welche Konsequenzen aus diesen Veränderungen resultieren
- welche Budgets für diese Vorhaben einzuplanen sind.

Zusätzlich wird in den Handlungsempfehlungen der Nutzen der jeweiligen Vorhaben dargestellt.

Bevor es auf Grundlage der Handlungsempfehlungen zu konkreten Entscheidungsvorlagen für die nächsten Budgetplanungsrounds kommt, sind die gegenseitigen Abhängigkeiten der einzelnen Maßnahmen zunächst zu analysieren, anschließend zu priorisieren und letztlich konkrete Projektvorhaben abzuleiten. Diese Aufgaben werden klassisch durch das Portfolio-Instrumentarium unterstützt. Die so ermittelten Projektvorschläge finden letztlich in den Budgetplanungsrounds Berücksichtigung und helfen durch deren Umsetzung sukzessiv die BI-Landschaft in die BI-Plattform zu überführen.

4 Gesammelte Erfahrungen und Ausblick

Die vorangegangenen Ausführungen stellen das Konzept und das Umfeld der BI-Bebauungsplanung im Rahmen der BI-Strategie des Volkswagen-Konzerns vor. So wird mit der BI-Bebauungsplanung ein Ansatz aufgezeigt, um die beschriebenen Ziele der BI-Strategie umzusetzen und mittelfristig zu erreichen.

Folgende zentrale Erfahrungen sind hierbei besonders herauszustellen:

- Die (BI-)Strategie schafft Orientierung. Eine verbindliche und ganzheitliche Umsetzung erfordert auch die Schaffung von strukturell und organisatorisch, geänderten Rahmenbedingungen.
- Die machbaren Umsetzungsschritte richten sich nach den jeweils aktuell gültigen Rahmenbedingungen, z. B. nach Budgetierung und organisatorischen Gegebenheiten.
- Große, querschnittliche Veränderungen werden der Gesamt-IT-Leitung in Form von Projektinitiativen vorgeschlagen, damit eine verbindliche Umsetzungsplanung (Budgetierung, Benennung von Verantwortlichen, etc.) erreicht werden kann.
- Die geforderte Trennung von Gravitationszentren und BI-Services ist bereits technisch erfolgt. Die organisatorische Trennung kann jedoch erst nach Anpassung der organisatorischen Rahmenbedingungen erfolgen.
- Für die Konkretisierung der BI-Bebauung ist die Vernetzung der beteiligten Fachbereichs- und IT-Stellen die unabdingbare Voraussetzung. Dieses wurde durch die Etablierung querschnittlich besetzter und organisierter Gremien erreicht.
- Voraussetzung für die angestrebte Errichtung eines Enterprise Data Warehouse bzw. einer Enterprise Data Warehouse-Schicht ist eine verfügbare und unmittelbar nutzbare, technische Infrastruktur. Nur dann sind konkrete Umsetzungsschritte planbar.
- Bereits ohne eine detaillierte Definition und Ausgestaltung der verschiedenen Sichten der Soll-Bebauung konnten aus der Analyse der Ist-Bebauung eine Reihe von Maßnahmen abgeleitet werden, um die Effizienz und Effektivität der BI-Landschaft zu steigern.

Obwohl sich die Etablierung der BI-Bebauungsplanung durch die BI-Governance derzeit noch in einer frühen Phase befindet, deuten erste Ergebnisse, die sich aus Gesprächen mit den Stakeholdern und aus der Analyse des derzeitigen Standes der BI-Bebauung z. B. der Prozess-System-Datenlandkarte ableiten ließen, auf das insgesamt hohe Potenzial der BI-Bebauungsplanung hin. So konnten fachlich bereits Redundanzen und Lücken aufgedeckt und beseitigt werden, technisch Optimierungspotentiale für einzelne Systeme und organisatorisch ein Arbeitskreis zum verbesserten Informationsaustausch initiiert werden.

Trotzdem gibt es noch viel zu tun:

- Der Bebauungsplan wurde zunächst konzentriert für die zwei Kernprozesse KAP und PP erstellt. Die Erweiterung um die Kernprozesse SUP und SPK steht bevor, um zum Jahresende eine erste vollständige und prozessübergreifende Sicht auf die BI-Landschaft zu erhalten.
- Die Weiterentwicklung und Anpassung der BI-Bebauungsplanung wird nicht am Grünen Tisch entworfen, sondern erfolgt anhand der Erfahrungen aus der Anwendung und konkreter Anforderungen. Die genutzten Werkzeuge wie die BI-Bebauungspläne und die Portfolios werden in regelmäßigen Abständen überarbeitet und an die Bedürfnisse angepasst. So werden z. B. auch die verwendeten Visualisierungskonzepte mit der Zeit zunehmen und sich weiterentwickeln.
- Das Konzern-Bebauungsmanagement führt derzeit zur Unterstützung der MCP-Arbeit ein zentrales Werkzeug ein. Hier ist zu prüfen, inwieweit das eingesetzte Werkzeug zukünftig auch bei der Erstellung des BI-Bebauungsplans eingesetzt werden kann.
- Die im Rahmen der BI-Bebauung in einem ersten Schritt generisch definierten Datenklassen gilt es nun in weiteren Schritten durch das BI-Datenmanagement zu konkretisieren.
- Die BI-Bebauungsplanung ist eine fortlaufende Aufgabe. Daher sind periodische Gespräche zur Fortschreibung der BI-Bebauungsplanung ebenso wie eine retrospektive Betrachtung des Erreichten vorgesehen.

5 Literaturverzeichnis

- [GJFG07] Goeken, M.; Johannsen, W.; Fröhlich, M., Glasner, K.: Sichten der IT-Governance. In: IT-Governance. Zeitschrift des ISACA Germany Chapter e.V., No. 1, Vol 1 (2007).
- [Mue06] Mühleck, K.: Innovative Geschäftsprozesse – Wertbeitrag eines modernen IT-Managements. In (Dietrich, L.; Schirra, W., Hrsg.): Innovationen durch IT – Erfolgsbeispiele aus der Praxis Produkte – Prozesse -. Springer-Verlag Berlin Heidelberg, 2006; S: 173 – 188.
- [LMW05] Lankes, J.; Matthes, F.; Wittenburg, A.: Softwarekartographie: Systematische Darstellung von Anwendungslandschaften. In: 7. internationale Tagung Wirtschaftsinformatik 2005, Bamberg, 2005.
- [SO07] Schönherr, M.; Offermann, P.: Unternehmensarchitektur als integrierende Sicht. In (Krallmann, H., Schönherr, M., Trier, M., (Hrsg.)): Systemanalyse im Unternehmen. Oldenbourg, München, 5., vollständig überarbeitete Auflage 2007; S. 29-56.
- [So08] Sommer, Th. et.al.: Business Intelligence-Strategie bei der Volkswagen AG. In (Dinter, B; Winter R., (Hrsg.)): Integrierte Informationslogistik. Springer-Verlag Berlin Heidelberg, 2008; S: 261 – 283.
- [TZ06] Trost, U.; Zirkel, M.: Wege aus dem Informationschaos. In: BI-Spektrum 1 (2006) 3, S. 11-16.

- [VW07] BI-Strategie des Volkswagen-Konzerns, Volkswagen AG, Wolfsburg, 2007.
- [Wi08] Winter, R. et.al.: Das St. Galler Konzept der Informationslogistik. In (Dinter, B; Winter R., (Hrsg.)): Integrierte Informationslogistik. Springer-Verlag Berlin Heidelberg, 2008; S: 1 – 16.

Intelligente Wohnungsmarktanalyse auf Basis einer Business Intelligence-Anwendung

Anforderungen, Modellierung und prototypische Realisierung

Roland Gabriel, Tobias Hoppe, Alexander Pastwa

Lehrstuhl für Wirtschaftsinformatik

Ruhr-Universität Bochum

Universitätsstraße 150

44780 Bochum

{rgabriel | thoppe | apastwa}@winf.rub.de

Abstract: Der Gegenstandsbereich des vorliegenden Beitrags richtet sich auf die Darstellung eines Ansatzes zur Gestaltung einer Business Intelligence-Anwendung (BI-Anwendung), die zur Entscheidungsunterstützung bei der Wohnungsmarktanalyse zum Einsatz kommen kann. Der im Folgenden zu präsentierende BI-Prototyp ist das Ergebnis eines Forschungsprojekts, das in Zusammenarbeit mit dem Institut für Wohnungswesen, Immobilienwirtschaft, Stadt- und Regionalentwicklung GmbH Bochum (InWIS) durchgeführt wurde. Die Wohnungsmarktanalyse eignet sich besonders dadurch als Einsatzfeld einer BI-Anwendung, weil Online-Marktplätze wie z. B. ImmobilienScout24 bereits die notwendige Datenbasis für eine derartige Analyse vorhalten. Eine wesentliche Triebfeder, die in diesem Zusammenhang den Einsatz einer BI-Anwendung zur Wohnungsmarktanalyse rechtfertigt, ist in der bisher fehlenden Möglichkeit begründet, eine verlässliche und fundierte Einschätzung der Potenziale des durch Intransparenz gekennzeichneten Wohnungsmarkts IT-gestützt zu ermöglichen. Um die an die BI-Anwendung formulierten Anforderungen zu erfüllen, wurden als Kerntätigkeiten im Entwicklungsprozess die Modellierung einer geeigneten multidimensionalen Datenstruktur und die Gestaltung aussagekräftiger, grafischer Reports, die eine schnelle Informationsverarbeitung gewährleisten, definiert. Im Ergebnis ließ sich ein BI-Prototyp implementieren, der sowohl für die Zwecke der Marktanalyse als auch zur Bestandsanalyse eingesetzt werden kann.

1 Wohnungsmarktanalysen als Einsatzfeld von Business Intelligence

Ein zentrales Charakteristikum des Wohnungsmarkts in Deutschland ist in seiner stark heterogenen Marktstruktur zu finden. Die Marktteilnehmer repräsentieren eine nahezu unüberschaubare Anzahl von Institutionen, Unternehmen sowie auch Privatpersonen, die hauptsächlich über klassische Anzeigen in lokalen Zeitungen miteinander kommunizieren und auf diese Weise eine Fülle von Informationen produzieren und austauschen. Insbesondere für Privatanbieter, aber auch für Unternehmen, die Immobilien zum Kauf oder zur Miete inserieren, führt die Intransparenz des Markts zur Schwierigkeit, fundier-

te Entscheidungen auf Basis verlässlicher Daten zu treffen. Mit dem Internet hat sich in jüngster Zeit ein weiteres Kommunikationsmedium zum Austausch von Informationen etabliert. Anbieter profitieren davon, ihre Angebote im Internet publizieren und ortsunabhängig vermarkten zu können, während die Nachfrager primär die kostenfreien und umfangreichen Suchmöglichkeiten in Anspruch nehmen, die weit über die Möglichkeiten klassischer Anzeigerecherchen hinausgehen.

Eine besondere Bedeutung für die internetgestützte Kommunikation zwischen anbietenden und nachfragenden Marktparteien nehmen Online-Marktplätze, wie zum Beispiel ImmobilienScout24, ein. Mit monatlich über 2,5 Millionen Besuchern [Ni07], rund 600 Millionen Seitenaufrufen, über 110 Millionen abgerufenen Exposés und über 70.000 Immobilienanbietern ist ImmobilienScout24 der größte Anbieter in diesem Markt. Obwohl nur etwa 50% der Marktteilnehmer ihre Angebote über Online-Marktplätze vermarkten, repräsentiert ImmobilienScout24 dennoch einen ausreichend großen Querschnitt über den Gesamtmarkt und eröffnet mit seinen quartalsweise gesammelten Datenbeständen neue Perspektiven zur Analyse des Wohnungsmarkts.

Um eine intelligente Wohnungsmarktanalyse IT-gestützt durchzuführen, war es erforderlich, ein Informationssystem zu gestalten, das eine flexible und performante Analyse des von ImmobilienScout24 bereitgestellten Datenbestands ermöglicht. Diese Aufgabe lässt sich mit Hilfe von Business Intelligence-Systemen (BI-Systeme) bewältigen. In einem integrierten Begriffverständnis lässt sich BI als Gesamtheit aller Werkzeuge und Anwendungen mit entscheidungsunterstützendem Charakter verstehen, die zur besseren Einsicht in das eigene Geschäft und damit zum besseren Verständnis in die Mechanismen relevanter Wirkungsketten führen [GGD07; KMU04]. Mit Blick auf diese Abgrenzung des BI-Begriffs steht dieser für eine begriffliche Klammer, die eine Vielzahl unterschiedlicher Ansätze, Werkzeuge sowie Anwendungen zur Analyse geschäftsrelevanter Daten zu bündeln versucht. Zu diesen gehört beispielsweise der Einsatz eines Data Warehouse (DW) [In96], das die zentrale Datenbasis für die Implementierung eines BI-Systems repräsentiert [GI01].

Dieser Beitrag erläutert die Ergebnisse eines Business Intelligence-Projekts, das in Zusammenarbeit mit dem Institut für Wohnungswesen, Immobilienwirtschaft, Stadt- und Regionalentwicklung GmbH Bochum (InWIS) durchgeführt wurde. Das InWIS ist ein gemeinnütziges Forschungs- und Transferinstitut. Zu den Tätigkeitsschwerpunkten des InWIS zählen sowohl die Durchführung grundlagenbezogener Forschung im Bereich des Wohnungswesens, der Immobilienwirtschaft sowie regionalwissenschaftlicher Fragestellungen als auch der Aufbau und die Pflege eines Informations- und Dokumentationszentrums. Als weiterer Aufgabenbereich des Instituts ist der Transfer wissenschaftlicher Erkenntnisse in Lehre und Praxis zu nennen. [In08] Gegenstand des Projekts war die Gestaltung eines BI-Systems zur Analyse des Wohnungsmarkts in Deutschland auf Basis eines Datenbestands, der von ImmobilienScout24 bereitgestellt wurde. Bei der Gestaltung der BI-Anwendung wurde als Datenbanksystem der Microsoft SQL Server 2000 verwendet, während die Implementierung des Data Warehouse mit Hilfe der Microsoft SQL Server 2000 Analysis Services erfolgte. Zur Erstellung und webgestützten Präsentation der Berichte kam die BI-Suite dynaSight von Arcplan zum Einsatz.

Der vorliegende Beitrag gliedert sich in sechs Abschnitte. Nach der Einleitung werden in Abschnitt 2 zunächst die Grundlagen zur Immobilienanalyse gelegt. Ausgehend von den Anforderungen an das BI-System, die in Abschnitt 3 den Gegenstand der Betrachtung bilden, widmet sich der darauf folgende Abschnitt der Modellierung des multidimensionalen Datenschemas als zentrales Element eines BI-Systems. Der daran anschließende Abschnitt 5 fokussiert auf die Präsentation des BI-Prototypen, bevor im letzten Abschnitt die Ergebnisse kritisch gewürdigt und weitere Ansatzpunkte für zukünftige Entwicklungen aufgezeigt werden.

2 Einführung in die Immobilienanalyse

Wie bei anderen Gütermärkten unterliegt auch der Immobilienmarkt zyklischen Entwicklungen. Eine auf die Immobilien ausgerichtete Analyse hat die Funktion, die Erwartungsbildung in Bezug auf Risiko und Ertrag des einzelnen Objekts zu unterstützen. Abgesehen von zufälligen Umwelteinflüssen sind die meisten Einflussfaktoren auf die Renditeentwicklung des Einzelobjekts einer quantitativen Analyse zugänglich, da sie auf Standorteigenschaften, lokalen Marktgegebenheiten, Eigenschaften der Immobilie selbst sowie der konkreten vertraglichen Ausgestaltungen beruhen [Ba99].

Immobilienanalysen lassen sich in Basis- und Spezialanalysen separieren. Während Basisanalysen der Erforschung der grundsätzlichen Potenziale und Probleme einer Liegenschaft dienen, sind Spezialanalysen auf den enger gefassten Informationsbedarf einzelner Entscheidungssituationen zugeschnitten [IV00]. Zu den Basisanalysen gehören zum Beispiel Standortanalysen, Marktanalysen, Gebäudeanalysen und Mietanalysen. Unter Spezialanalysen werden beispielsweise Bewertungsanalysen, Investitionsanalysen, Finanzanalysen und Machbarkeitsstudien zusammengefasst. Im Rahmen des BI-Projekts lag der Fokus auf der Standort- sowie Marktanalyse, da diese oftmals als Grundlage für zusätzliche Spezialanalysen dienen.

Eine Standortanalyse als Ausprägung einer Basisanalyse setzt sich mit den räumlichen Rahmenbedingungen eines Objekts auseinander und wird auf Makroebene (Stadt, Umland) und Mikroebene (Grundstück) durchgeführt [Mu96]. Für diese Unterteilung ist auch der zeitliche Aspekt von Bedeutung. Die Elemente auf der Makroebene zeichnen sich in der Regel durch einen relativ gleichmäßigen Entwicklungsverlauf aus, was beispielsweise die Wirtschaftsstruktur oder die Verkehrsstruktur anbetrifft, oder sie bleiben sogar langfristig konstant, z. B. mit Blick auf die Entfernung zu den Nachbarstädten oder die Zentralörtlichkeit. Die Mikrofaktoren können beispielsweise in Bezug auf die Grundstücksgröße zwar ebenfalls eine Konstanz aufweisen, sich aber auch innerhalb kurzer Zeit in ihrer Ausprägung oder Bedeutung ändern (Sozialstruktur, Prestige der Lage).

Eine Marktanalyse als eine weitere Unterkategorie der Basisanalyse führt eine kurz- und mittelfristige Abschätzung des Vermietungs- und Verkaufserfolgs einer Nutzungskonzeption für ein Objekt durch. Dazu werden das Angebot und die Nachfrage auf dem Immobilienmarkt untersucht. Eine Angebotsanalyse erfasst den Flächenbestand, der nach verschiedenen Kriterien und Merkmalen segmentiert wird. Die Nachfrageanalyse

beschäftigt sich mit dem aktuellen Flächenbedarf, der vornehmlich durch vorhandene Gesuche abgeleitet wird, aber auch anhand der momentanen Absorptionsrate kurzfristig abgeschätzt werden kann. In der Preisanalyse werden differenzierte Entwicklungen von Miet- und Kaufpreisen ausgewertet.

Die Nachfrageanalyse, die auf volkswirtschaftlichen Kennzahlen basiert, untersucht oftmals auch die demografische Situation oder die Veränderung des Einkommens der Bevölkerung, was sich auf die Verschiebung der Wohnungsnachfrage auswirkt. Eine derartige Analyse wurde in dem BI-Projekt nicht betrachtet.

3 Anforderungen an das BI-System

Das Ziel des BI-Projekts lag darin, für das InWIS eine leicht zu bedienende Lösung zu gestalten, mit Hilfe derer sich auf eine flexible und intuitive Weise relevante Informationen des Wohnungsmarkts ermitteln lassen. Unter Einsatz des BI-Tools sollen Informationen als Entscheidungsgrundlage für wohnungswirtschaftliche Konzeptionen gewonnen werden. Zur Konkretisierung dieses Ziels ließen sich folgende Teilanforderungen identifizieren: Ein großer Interessenschwerpunkt des InWIS lag auf der Nutzung von Berichten, die nach dem Vorbild einer Ampelkarte eine Identifikation und Analyse markanter wohnungsmarktbezogener Kennzahlen innerhalb des Gebiets der Bundesrepublik Deutschland zulassen. Eine weitere Anforderung richtete sich auf die Implementierung einer Zeitreihenanalyse. Hierbei waren geeignete Dimensionen auszuwählen, die über einen bestimmten Zeitablauf signifikante Entwicklungen bei den Messwerten aufzeigen. Für das InWIS war darüber hinaus die Frage von Interesse, welche Eigenschaften Immobilien aufweisen sollen, um in bestimmten geografischen Gebieten einen größtmöglichen Gewinn bzw. eine zeitnahe Vermarktung zu gewährleisten. Da die Aussagequalität einer Analyse im Wesentlichen mit der Qualität und Quantität der verarbeiteten Inputdaten zusammenhängt, musste zur effizienten und effektiven Nutzung des BI-Tools eine Bereinigung des zur Auswertung vorliegenden, umfangreichen Datenmaterials vorgenommen werden. Das wichtigste Ziel der Datenbereinigung lag darin, die Kennzahlen in ihrem Wertebereich zu beschränken, um irrtümliche Werte herauszufiltern.

4 Multidimensionale Datenmodellierung

Eines der wesentlichen Erfolgsfaktoren bei der Gestaltung einer BI-Lösung ist die Formulierung eines geeigneten multidimensionalen Datenmodells. Folglich liegt beim Modellierungsprozess eine der zentralen Aufgaben in der Identifizierung relevanter Dimensionen und Messwerte bzw. Fakten (Measures) [KMU04]. Aufbauend auf einem relationalen Datenmodell, welches sich aus dem von ImmobilienScout24 zur Verfügung gestellten Datenmaterial ergab, wurde das multidimensionale Datenschema modelliert und als eigener Datenwürfel erstellt. Dabei sind im Hinblick einer verlässlichen Nutzung der in den verschiedenen Dimensionstabellen abgelegten Attribute und Werte beispielsweise im Rahmen eines OLAP oder Data Mining hohe Anforderungen an die Datenintegrität, Datenkonsistenz, Datensicherheit und Datenschutz zu erfüllen [MB98; GR95].

In der Regel gehen der Modellierung eines dem BI-System zugrunde liegenden multidimensionalen Datenschemas umfangreiche Tätigkeiten zur Datenaufbereitung voraus, die mit dem Ziel durchgeführt werden, den auszuwertenden Datenbestand in eine konsolidierte und konsistente Form zu überführen [Ba06; KMU04]. Vor dem Hintergrund der Tatsache, dass es den Benutzern von ImmobilienScout24 bei vielen Datenfeldern frei steht, beliebige und insbesondere auch unsinnige Eingaben zu den Angeboten sowie zu den Gesuchen zu tätigen, richtete sich der Aufwand insbesondere auf die Aufbereitung fehlerhafter sowie fehlender Einträge. Ein derartiger Datenzustand ist zum einen auf Absicht im Falle so genannter Scherzangebote zurückzuführen, andererseits Folge einfacher Tippfehler bei der Aufgabe des Inserats. Für die Aufbereitung waren zunächst klare Grenzen für den Filterprozess der untersuchungsrelevanten Kennziffern zu identifizieren. Um einerseits eine möglichst breite Datenbasis für die angestrebten Wohnungsmarktanalysen zu erhalten, andererseits auch genügend Ausreißer, die die Untersuchungen verfälschen würden, abzuschneiden, wurden jeweils die 0,5% niedrigsten und höchsten Werte einer Kennziffer nicht zur Auswertung herangezogen. Die aus dieser Vorgehensweise resultierenden gültigen Wertebereiche sind der folgenden Tabelle zu entnehmen:

stochastische Werte			
	Untergrenze	Obergrenze	Quantil α
Wohnfläche	8 qm	3.517 qm	1%
Kaltmiete	2 €	115.670 €	1%
Kaufpreis	55 €	11.250.000 €	1%
Nebenkosten	0,6 €	17.411 €	1%
Zimmerzahl	1	500	1%

Abbildung 1: Stochastische Grenzen der Datenbasis

Eine der aufbereitenden Maßnahmen richtete sich darauf, beim Kaufpreis alle Angebote von nur einem Euro auszuschließen, da diese in der Regel Versteigerungen darstellen, die nicht den realen Kaufpreis repräsentieren und auf diese Weise die Analysen verfälschen. Darüber hinaus erfolgte für alle Fakten die Filterung aller negativen Werte, ehe in einem zweiten Schritt zur Eliminierung von Datenausreißern die 0,5% größten und kleinsten Datensätze verworfen wurden. Diese Filterprozesse sind in Datensichten eingearbeitet, anstatt die vermeintlich fehlerhaften Dateneinträge physisch aus der Datenbank zu löschen. Die Verwendung von Sichten bietet in diesem Zusammenhang eine flexible Möglichkeit zur Kalibrierung des Verwerfungsbereichs für die Wertausprägungen der in der Tabelle aufgeführten Attribute. Das Ausgangsdatenmaterial bleibt damit unverändert, so dass alle im Projektverlauf durchgeführten Änderungen reversibel sind. Eine veränderte Datensicht lässt sich damit sofort und ohne strukturelle Anpassungen der mehrdimensionalen Datenmodellierung generieren.

Nach diesen vorbereitenden Maßnahmen konnten aus dem relationalen Datenmodell folgende Dimensionen ermittelt und in Form eines Snowflake-Schemas modelliert werden. Ein bei ImmobilienScout24 eingestelltes Immobilienangebot lässt sich in die vier

Hauptkategorien „Wohnen zur Miete“, „Wohnen nach Kauf“ sowie „Gewerbeobjekte“ und „Anlageobjekte“ zuordnen. Aus dem Datenbestand von ImmobilienScout24 wurden insgesamt sieben Dimensionstabellen samt Attributsausprägungen identifiziert und im Datenmodell hinterlegt. Die Dimensionstabellen beziehen sich auf die Heizungs-, Vermarktungs- und Nutzungsarten sowie den Immobilientyp, Objektkategorien, Objektzustände und die Orte. Jedes angebotene Objekt erhält eine eindeutige Objekt-ID und wird zusammen mit seinen charakterisierenden Attributsausprägungen in einer Faktentabelle gespeichert. Die Faktentabelle enthält neben einer eindeutigen Objekt-ID (für die angebotenen Immobilien) und den Sekundärschlüsseln, die die Primärschlüssel der den jeweiligen Objekten charakterisierenden Dimensionstabellen referenzieren, die auszuwertenden Fakten. Zur Abbildung der Zeit als siebte Dimension dieses Datenmodells wurde die Faktentabelle um eine Attributspalte für das Quartal erweitert. Auf diese Weise ließ sich jedem Datensatz in der Faktentabelle genau ein Quartalswert zuordnen. Dieser Schritt war nötig, um zeitliche Entwicklungen aufzeigen zu können, die bei der Wohnungsmarktanalyse bislang nur selten verfügbar sind.

4.1 Dimensionen

Immobilientyp

Die Dimension Immobilientyp dient der grundlegenden Kategorisierung der angebotenen Immobilienobjekte. Auf der obersten Aggregationsstufe werden die Objekte in Mietwohnungen, Miethäuser, Häuser zum Kauf bzw. Anlageobjekte eingeteilt. Unterkategorien schlüsseln diese Objekte weiter auf, so dass eine exakte Differenzierung, beispielsweise der Häuser in Reihenhäuser, Reihenendhäuser, Reiheneckhäuser, Appartementshäuser u. ä., möglich ist.

Zeit

Die Zeit-Dimension ermöglicht die Analyse der Fakten nach im Vorfeld zu definierenden Zeiträumen. In dem BI-System wurden die Hierarchiestufen „Jahr“ und „Quartal“ realisiert.

Kategoriedimension

Mit Hilfe dieser Dimension lassen sich aus dem Datenmaterial von ImmobilienScout24 künstliche Dimensionstabellen zusammenfassen, die dazu dienen, den Wohnungsmarkt bezüglich bestimmter Ausprägungen diverser Fakten zu segmentieren. Folgende Kennzahlen wurden durch eine Kategorisierung dimensioniert:

- Zimmeranzahl
- Kaufpreis
- Mietpreis
- Wohnfläche
- Grundfläche
- Etagenanzahl.

Vermarktungs- und Nutzungsart

Diese beiden Dimensionen geben an, ob ein Objekt vermietet, verpachtet oder verkauft werden soll und ob es für die weitere Nutzung als Miet-, Pacht- bzw. Kaufobjekt vorgesehen ist.

Objektzustand

Analysen zum qualitativen Zustand eines Immobilienobjekts lassen sich mit der Dimension Objektzustand durchführen. Diese Dimension erlaubt Auswertungen zur Frage, wie stark ein Marktsegment in Bewegung ist. Eine größere Anzahl angebotener Neubauten in gutem Zustand deutet beispielsweise auf einen verstärkten Markteintritt verschiedener Marktteilnehmer hin, wogegen ein lokales Segment, in dem lediglich bestehende Bau-substanz inseriert wird, ein Indiz für einen stagnierenden Markt sein kann.

Heizungsart

Die Heizungsart-Dimension gibt Auskunft über die eingesetzte Heizungsanlage im angebotenen Immobilienobjekt. Ausprägungen dieser Dimension können beispielsweise Zentralheizung (Öl/Gas), Etagenheizung, Fernwärme, Elektroheizung, Kohle, u. ä. sein.

Orte

Mit Hilfe der Ort-Dimension lassen sich die angebotenen Immobilienobjekte den entsprechenden Postleitzahlen zuordnen. Eine Hierarchisierung innerhalb dieser Dimension erfolgt über die Ebenen „Gemeinde/Stadt“, „Landkreis/Kreisfreie Stadt“, „Bundesland“ und „Land“.

4.2 Messwerte (Measures)

Um eine Wohnungsmarktanalyse durchführen zu können, ließen sich folgende Messwerte identifizieren bzw. nachträglich in der Faktentabelle erstellen. Die Metriken unterteilen sich in drei grundlegende Gruppen. Numerische Charakteristika, wie etwa der Kaufpreis oder die Kaltmiete sind Zahlenwerte, die direkt für Analysezwecke verwendet werden können. Ausstattungsmerkmale, wie beispielsweise das Vorhandensein eines Balkons oder Gartens sind im Sinne ihrer Ja-/Nein-Ausprägungen binär 0/1-kodiert, um effiziente Analysen auf dem umfangreichen Datenbestand durchführen zu können. Nachfragekennziffern als letzte Kennzahlengruppe versuchen schließlich, Indizien für eine bestehende Nachfrage nach den Immobilienobjekten aus den Angeboten abzubilden.

Die Speicherung der Nachfragekennziffern erfolgt automatisch und anonymisiert. Während Weiterempfehlungen und Kontaktaufnahmen auf den ersten Blick geeignet erscheinen, die Nachfrage für ein Objekt zu quantifizieren, ist gleichzeitig zu konstatieren, dass das Internetportal keinesfalls die einzige Möglichkeit zur Kontaktaufnahme zwischen den Kunden und den Anbietern darstellt. Alternativ kann ein Kontakt telefonisch oder per E-Mail hergestellt werden, wodurch tatsächlich existierende Nachfrage in diesem Kontext unbeobachtet bleibt. Die Kennzahl „Hits“ repräsentiert eine umfassende Kennziffer für die Nachfrage, die angebotsseitig gespeichert ist und informiert, wie häufig die Benutzer dieses Objekt betrachtet haben. In diesem Zusammenhang besteht das Problem, dass diese Kennziffer keine konkrete Aussage über einen Geschäftsabschluss liefert. Anhand der Daten lässt sich folglich nicht feststellen, ob ein Geschäft tatsächlich zu-

stande gekommen ist. Einen weiteren Hinweis für einen Geschäftsabschluss bietet das Measure „Laufzeittage“, welches für jedes Objekt registriert, wie lange es inseriert wird, bevor der Anbieter es zurückzieht oder es ausläuft. Dieser Wert ist ebenfalls nur ein Indiz für einen stattgefundenen Geschäftsabschluss, da ein Inserat aufgrund vollständig fehlender Nachfrage zurückgezogen werden kann. Ferner ist der Fall denkbar, dass auf andere Portale oder Vermarktungsformen ausgewichen wird, um nicht länger die monatlichen Inseratskosten des ImmobilienScout24-Portals tragen zu müssen.

Im Folgenden sind weitere wichtige Messwerte aufgeführt, die sich aus den oben genannten berechnen.

Quadratmeterpreise Miet/Kauf:

Das Faktum stellt den Quotient aus Kaltmiete bzw. Kaufpreis und Wohnfläche dar. Durch diese Verhältniszahl können die Miet-/Kaufpreise verschiedener Objekte miteinander verglichen werden.

Hits pro Tag/Objekt:

In dieser Messgröße wird die Häufigkeit des Auftretens eines Hits pro Laufzeittag bzw. Objekt festgehalten. Auf diese Weise lässt sich die Nachfrage nach einzelnen Objekten quantifizieren.

Durchschnittliche Quadratmeterpreise/Wohnfläche:

Um regionale bzw. lokale Vergleichswerte zu erhalten, sind, abhängig von der gewählten regionalen Aggregationsstufe, die in diesem Segment durchschnittlichen Quadratmeterpreise bzw. Wohnflächen als Kennzahl hinterlegt.

Durchschnittliche Laufzeittage:

Diese Kennzahl gibt Aufschluss darüber, ob sich die Objekte verhältnismäßig schnell bzw. schwierig vermarkten lassen.

5 Der BI-Prototyp

Der webfähige BI-Prototyp bietet dem Anwender über ein Portal Zugriff auf die angelegten Standard-Reports. Im Vordergrund der Analysen stehen dabei folgende beispielhaft für einen Bauträger interessante Fragestellungen:

- Ist ein Standort generell für die Entwicklung bestimmter Wohnungstypen geeignet?
- Welche Ausstattungsvarianten sollte das Gebäude aufweisen?
- Wie groß sollten die Wohneinheiten konzipiert werden?
- Wie viele Räume werden am ehesten nachgefragt?
- Wie hoch sollte der Quadratmeterpreis der Wohnungen/Häuser kalkuliert werden, um einerseits einen größtmöglichen Gewinn, andererseits eine zeitnahe Vermarktung zu gewährleisten?

- Was sind alternative Immobilientypen, die sich für einen bestimmten Standort besser eignen?

Wie in Abbildung 2 verdeutlicht ist, werden alle immobilienwirtschaftlich relevanten Informationen in Form verschiedener Berichte präsentiert, die einer der beiden Berichtsklassen („Marktanalysen“ oder „Bestandsanalysen“) zugeordnet sind. Die vom InWIS am häufigsten verwendeten Berichte sind Gegenstand der folgenden Ausführungen.

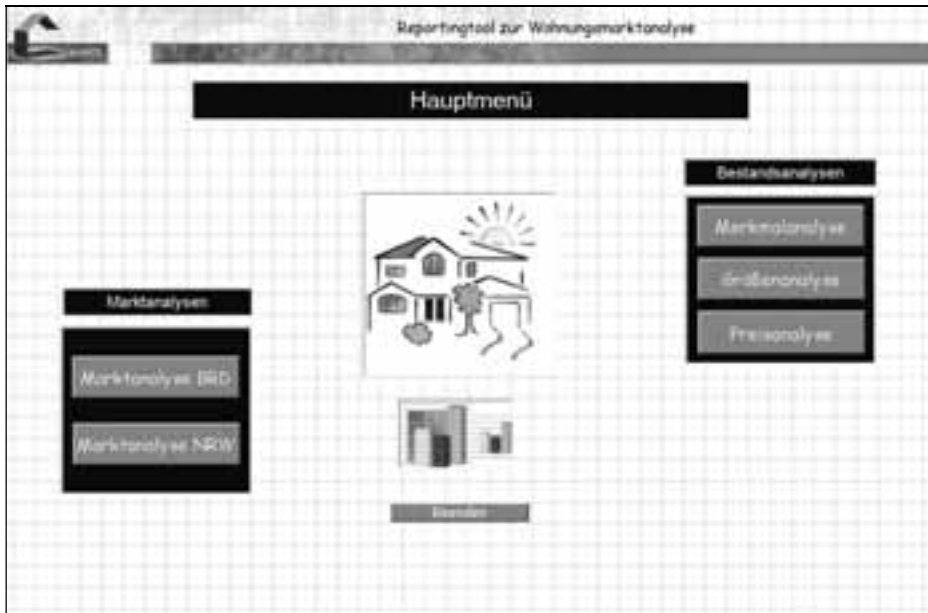


Abbildung 2: Portal des Reportingtools zur Wohnungsmarktanalyse

Um einen Tradeoff zwischen Informationsüberflutung und Übersichtlichkeit zu vermeiden, sind die Reports größtenteils grafisch aufbereitet. Die Marktanalysen für die Bundesrepublik Deutschland bzw. für Nordrhein-Westfalen visualisieren in Form einer Landkarte mit integrierter Ampelfunktionalität regionale Unterschiede auf verschiedenen Aggregationsniveaus. Dabei können in Auswahlmenüs die Metriken selektiert und die Dimensionselemente eingeschränkt werden, die in dem Report grafisch angezeigt werden sollen. Diese aktiven Landkarten dienen als Ausgangspunkt für weitere Detailanalysen. Dargestellt werden, wie in der Abbildung 3 exemplarisch visualisiert, die durchschnittlichen Quadratmeterpreise aller zum Kauf angebotenen Immobilien für die einzelnen Bundesländer. Die farbliche Segmentierung spiegelt die Preisdifferenzen wider. Gelbe Sektoren markieren die Bundesländer, die innerhalb eines 20-prozentigen Durchschnittsintervalls liegen. Selbstverständlich kann die Größe des Durchschnittsintervalls in den Einstellungen den individuellen Bedürfnissen angepasst werden. Die oberhalb dieser Marke liegenden werden grün, die unterhalb rot dargestellt. Ausgehend hiervon lässt sich durch einen Klick auf das entsprechende Bundesland ein Drill-Down auf die untergeordnete Ebene der geografischen Hierarchie (Landkreise eines Bundeslandes) ausführen, die ebenfalls zu einer grafischen Auswertung der zuvor getroffenen Auswahlkriterien führt.

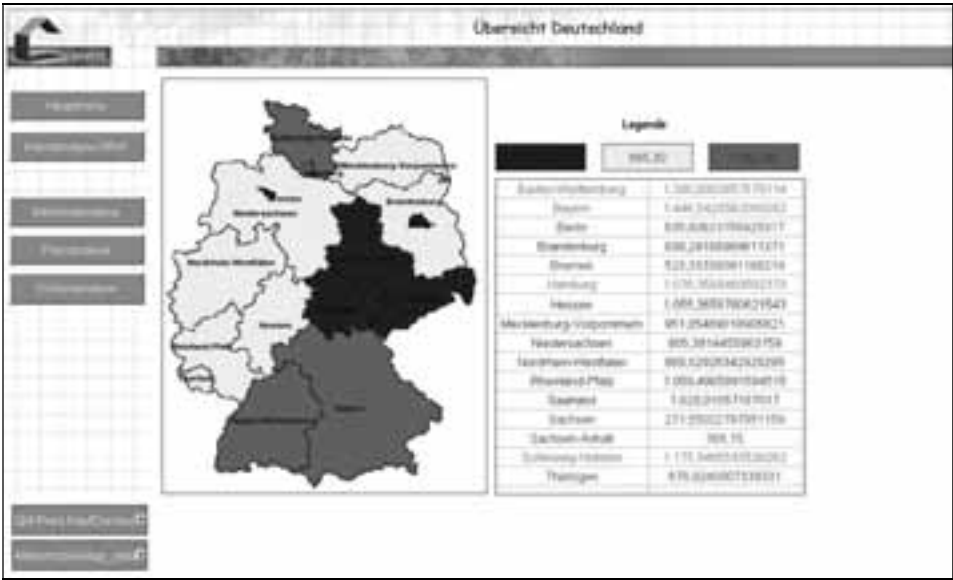


Abbildung 3: Marktanalyse Bundesrepublik Deutschland

Die Bestandsanalysen untersuchen den Immobilienbestand nach unterschiedlichen Kriterien. In dieser Kategorie lassen sich drei Reports aufrufen, die erstens die Objektmerkmale analysieren, zweitens die Größe der Immobilien auswerten und drittens eine Preisanalyse der Objekte ermöglichen.

Unter allen Reports stellt die Merkmalanalyse den Bericht mit der größten Informationsdichte dar. In diesem Report können alle Ausstattungsmerkmale der angebotenen Immobilien flexibel untersucht werden. Hierbei erfolgt sowohl eine grafische als auch eine tabellarische Darstellung der ausgewählten Fakten wie beispielsweise die Hits pro Tag bzw. pro Objekt oder die durchschnittlichen Laufzeittage. Die einzelnen Ausstattungsmerkmale, kombiniert mit den entsprechenden Kennzahlen, ermöglichen somit einen Überblick, wie Immobilien in bestimmten Wohnregionen konzipiert werden müssen, um einen optimalen Absatz dieser zu erlangen.

Weitere wichtige Fragestellungen – gerade im Zusammenhang mit der Erschließung neuer Bebauungsgebiete – richten sich auf die Größe der Grundstücke und Wohnungen. Die Größenanalyse stellt zu dieser Frage die notwendigen Informationen zur Verfügung. Der entsprechende Report ermöglicht durch die Auswertung der Hits auf bestimmte Objekte die Nachfrage – klassifiziert in Größenintervalle, nach Wohnungen, Häusern sowie Grundstücken – in der gewählten Region abzubilden. Aus diesen Informationen lassen sich Schlüsse bezüglich der Gestaltung der Grundstücksgrößen in Neubaugebieten bzw. der Wohnungsgrößen bei Eigentumswohnungen ziehen. Die Abbildung 4 veranschaulicht exemplarisch einen derartigen Bericht für die Nachfrage nach Reihenhäusern, die nach Wohnungsgrößen in m² klassifiziert sind. Der Betrachtungszeitraum umfasst die Jahre 2005 und 2006. Als Measures sind für die jeweilige Wohnungsgröße die Anzahl der angebotenen Reihenhäuser sowie die entsprechenden Hits pro Tag ausgewählt.

Letztendlich spielt der erzielbare Marktpreis beim Verkauf von Grundstücken, Häusern und Wohnungen für den Erfolg des Bauträgers eine wesentliche Rolle. Der Bericht „Preisanalyse“ untersucht die in der Vergangenheit veräußerten Objekte, wobei eine gezielte Eingrenzung des Untersuchungsobjekts bezüglich der Ausstattungsmerkmale und dem Immobilientyp vorgenommen werden muss. Hierdurch lassen sich für konkrete Bauvorhaben die optimalen Verkaufspreise ermitteln, die in einem bestimmten Wohngebiet einer Stadt erzielbar sind. Derartige Preisanalysen lassen sich selbstverständlich auch für Mietobjekte durchführen.

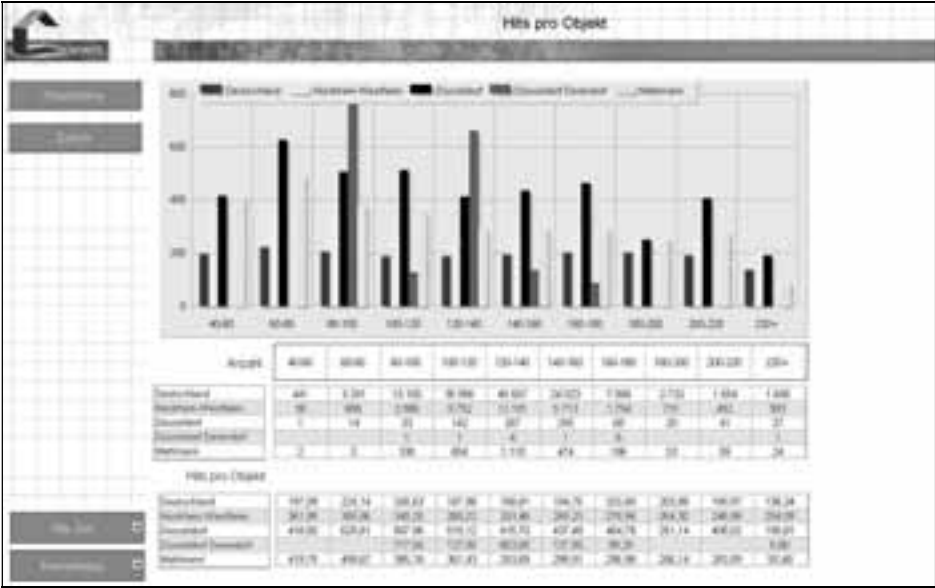


Abbildung 4: Größenanalyse für Reihenhäuser

6 Zusammenfassung und Ausblick

Das entwickelte BI-Tool zur Analyse des Wohnungsmarkts verdeutlicht, dass der Einsatzbereich von BI-Anwendungen nicht ausschließlich auf die Entscheidungsunterstützung in klassischen betrieblichen Funktionen wie Marketing, Controlling oder Produktion begrenzt ist. Ebenso können „exotische“ Einsatzzwecke wie die Wohnungsmarktanalyse mit Hilfe solcher Systeme sinnvoll unterstützt werden. Der auf der Basis von Microsoft SQL Server und Arcplan dynaSight entwickelte BI-Prototyp zeigt große Potenziale für umfangreiche, datenintensive Marktanalysen auf, die in einer derartigen Datenqualität und Flexibilität den Marktteilnehmern bisher noch nicht oder lediglich sehr eingeschränkt zur Verfügung stehen, um den Markt für ihre Zwecke adäquat zu segmentieren. Da zu erwarten ist, dass die über Online-Marktplätze abgewickelten Transaktionen in den kommenden Jahren weiter zunehmen werden, ist in Aussicht gestellt, dass der hier vorgestellte Ansatz in Zukunft noch von einer wesentlich breiteren Datenbasis profitieren und folglich an Aussagefähigkeit weiter gewinnen wird.

Sowohl die Markt- als auch die Bestandsanalysen, die sich mit Hilfe des BI-Tools durchführen lassen, leisten einen Beitrag dazu, den Wohnungsmarkt mit seinen umfangreichen und komplexen Verflechtungen besser zu verstehen. Die Marktanalysen geben einen guten visuellen Überblick der marktrelevanten Kennzahlen. Die Bestandsanalysen erhöhen diesen Informationsgewinn durch detaillierte Auswertungen auf einem sehr geringen Aggregationsniveau. Somit lassen sich auf bestehende oder neue Wohngebiete abgestimmte Bauvorhaben planen, die sich zum einen optimal in das lokale geografische Gefüge einordnen, aber auch den Bedürfnissen und Ansprüchen der Käufer-/Mieterklientel entsprechen.

7 Literaturverzeichnis

- [Ba06] Bange, Carsten: Werkzeuge für analytische Informationssysteme. In (Chamoni, P.; Gluchowski, P., Hrsg.): Analytische Informationssysteme – Business Intelligence-Technologien und -Anwendungen. Springer-Verlag, 3. Aufl., Berlin/Heidelberg, 2006; S. 89-110.
- [Ba99] Baum, Andrew: Evidence of cycles in European commercial real estate markets – and some hypotheses. <http://www.reading.ac.uk/LM/LM/fulltxt/0500.pdf>, 1999.
- [GGD07] Gluchowski, Peter; Gabriel, Roland; Dittmar, Carsten: Management Support Systeme und Business Intelligence – Computergestützte Informationssysteme für Fach- und Führungskräfte. Springer-Verlag, 2. Aufl., Berlin/Heidelberg, 2007.
- [Gl06] Gluchowski, Peter: Business Intelligence – Konzepte, Technologien und Einsatzbereiche. In (Hildebrand, K., Hrsg.): HMD (Praxis der Wirtschaftsinformatik). dpunkt.verlag, Heft 222, 2001; S. 5-15.
- [GR95] Gabriel, Roland; Röhrs, Heinz-Peter: Datenbanksysteme – Konzeptionelle Datenmodellierung und Datenbankarchitekturen. Springer-Verlag, 2. Aufl., Berlin/Heidelberg, 1995.
- [In96] Inmon, William H.: Building the Data Warehouse. John Wiley & Sons, 2. Aufl., New York et al., 1996.
- [In08] InWIS: Institutsprofil. http://www.inwis.de/htm/institut/inst_ges.htm, 2008.
- [IV00] Isenhöfer, Björn; Väth, Arno: Immobilienanalyse. In (Schulte, K.-W., Hrsg.): Immobilienökonomie, Band 1, Betriebswirtschaftliche Grundlagen. Oldenbourg-Verlag, 2. Aufl., München, 2000; S. 319-379.
- [KMU04] Kemper, Hans-Georg; Mehanna, Walid; Unger, Carsten: Business Intelligence – Grundlagen und praktische Anwendungen. Vieweg-Verlag, Wiesbaden, 2004.
- [MB98] Mucksch, Harry; Behme, Wolfgang: Das Data Warehouse-Konzept als Basis einer unternehmensweiten Informationslogistik. In (Mucksch, H.; Behme, W., Hrsg.): Das Data Warehouse-Konzept. Gabler-Verlag, 3. Aufl., Wiesbaden, 1998; S. 33-100.
- [Mu96] Muncke, Günter: Standort- und Marktanalyse in der Immobilienwirtschaft – Ziele, Gegenstand, methodische Grundlagen, Datenbasis und Informationslücken. In (Schulte, K.-W., Hrsg.): Handbuch Immobilien-Projektentwicklung. Immobilien Informationsverlag Rudolf Müller, Köln, 1996.
- [Ni07] NielsenNetRatings: Über uns – Aktuelle Kennzahlen. <http://www.immobilienscout24.de>, 2007.

Ein mandantenfähiges Data-Warehouse-System für die bayerischen Fachhochschulen - Anforderungen, Konzeption und Realisierung

Tim-Oliver T. Förtsch, Stefan Hartmann, Carsten Jürck, Julia Ringler

Wissenschaftliches Institut für Hochschulsoftware
der Universität Bamberg (ihb)
Feldkirchenstr. 21
96045 Bamberg

{tim.foertsch | stefan.hartmann | carsten.juerck | julia.ringler}@ihb.uni-bamberg.de

Abstract: Seit einigen Jahren befindet sich das deutsche Hochschulwesen im Umbruch, was sich aktuell in der Einführung der leistungsorientierten Mittelverteilung oder der Studienbeiträge zeigt. Der Wandel in Richtung New Public Management führt zu einem erhöhten Bedarf an fundierten Informationen über die eigenen Hochschulprozesse. Zur Unterstützung von Entscheidungen haben sich mittlerweile auch im Hochschulumfeld Data-Warehouse-Systeme etabliert. Um dieser Entwicklung gerecht zu werden, wurde im Rahmen des Projekts CEUS ein mandantenfähiges Data-Warehouse-System für die bayerischen Fachhochschulen realisiert. Ziel des neu konzipierten Systems war es, eine Basis für eine standardisierte, flexible Informationsversorgung bei einheitlicher Begriffsdefinition für alle Fachhochschulen zu schaffen. Das entwickelte mandantenfähige Data-Warehouse-System stellt hierfür eine geeignete Lösung dar.

1 Einleitung

Der Bedarf an gestaltungs- und steuerungsrelevanten Informationen ist in den letzten Jahren im Hochschulumfeld stark angestiegen. Hierfür sind u. a. die Festlegung von Zielvereinbarungen und die Überwachung der Zielerreichung oder die leistungs- und belastungsorientierte Mittelverteilung verantwortlich (New Public Management; vgl. z. B. [Ku08]). Eine wesentliche Voraussetzung zur Bewältigung dieser Herausforderungen stellt die Verfügbarkeit von validen und vergleichbaren Daten für das Hochschulmanagement dar.

Das Bayerische Staatsministerium für Wissenschaft, Forschung und Kunst (StMWFK) fördert mit dem Projekt CEUS¹ (Computerbasiertes Entscheidungsunterstützungssystem für die Hochschulen in Bayern) die Entwicklung und den Betrieb von Data-Warehouse-

¹ <http://www.ceushb.de>.

Systemen (DWH-Systemen) zur umfassenden Informationsversorgung aller Entscheidungsbereiche im bayerischen Hochschulwesen. Das wissenschaftliche Institut für Hochschulsoftware der Universität Bamberg (ihb²) ist seit 2003 mit der Einführung und Weiterentwicklung von CEUS am StMWFK und an den staatlichen Hochschulen beauftragt. Das ihb versteht sich hierbei als Kompetenzzentrum für die Entwicklung und Einführung von Softwarelösungen zur Unterstützung von Hochschulprozessen. Zum Leistungsspektrum des ihb gehören ebenfalls Beratungsdienstleistungen und Schulungsmaßnahmen.

Eine grundlegende Erweiterung erfährt das Projekt CEUS mit der im Jahr 2006 begonnenen Einführung an den 17 bayerischen Fachhochschulen. Dabei galt es einen Lösungsansatz zu entwickeln, der allen Fachhochschulen jeweils ein lokales DWH-System offeriert, das zugleich in die bestehende CEUS-Gesamtarchitektur eingebunden ist. Wichtige Ziele waren dabei auf Seiten des ihb, den Implementierungs- und Wartungsaufwand über alle lokalen DWH-Systeme so weit wie möglich zu reduzieren sowie auf Seiten der Fachhochschulen die Vergleichbarkeit des Berichtswesens untereinander und – im Hinblick auf die Zielvereinbarungen – mit dem StMWFK sicherzustellen.

Ziel des vorliegenden Beitrags ist es, die Anforderungsanalyse, Konzeption und Realisierung eines mandantenfähigen DWH-Systems für die CEUS-Einführung an den bayerischen Fachhochschulen aufzuzeigen. Hierzu werden in Kapitel 2 Anforderungen an die Architektur eines solchen Systems aufgezeigt. Kapitel 3 umfasst die Konzeption und Kapitel 4 darauf aufbauend die Realisierung des DWH-Systems, welches den beschriebenen Anforderungen genügt. In Kapitel 5 folgen eine Zusammenfassung sowie ein kurzer Erfahrungsbericht über den mit dieser Architektur erzielten Nutzen.

2 Anforderungen an ein DWH-System für die Fachhochschulen

Ausgangspunkt einer Analyse der Anforderungen an ein DWH-System für die bayerischen Fachhochschulen ist die Einbindung der einzelnen Systeme in die CEUS-Gesamtarchitektur. Diese besteht aus einer hierarchisch verteilten Data-Warehouse-Architektur mit autonomen Teilsystemen auf Hochschul- und auf Landesebene (vgl. [Si01]). Die lokalen DWH-Systeme der bayerischen Hochschulen basieren auf Detaildaten zu den Domänen *Studierende und Prüfungen*, *Finanzmittel* sowie *Personal und Stellen*. Das Data-Warehouse auf Landesebene beinhaltet aggregierte, konsolidierte und damit vergleichbare Informationen über alle Hochschulen des Landes. Mit dieser Architektur ist es möglich, sowohl den Informationsbedarf einzelner Hochschulen (unter Berücksichtigung derer individuellen Organisations- und Fächerstrukturen) als auch den Bedarf des StMWFK an hochschulübergreifenden Vergleichswerten zu berücksichtigen. Die DWH-Systeme stellen hiermit eine Dienstleistung für die Hochschuleinrichtungen (z. B. Studiendekane, Fakultäten) bei gleichzeitiger Entlastung der Verwaltung zur Verfügung.

² <http://www.ihb.uni-bamberg.de>.

Rückblickend betrachtet führte die sukzessive CEUS-Einführung an den bayerischen Universitäten zu einer Diversifikation der lokalen DWH-Systeme. Von den Universitäten wurde zu Projektbeginn ein eigenständiger Betrieb der lokalen DWH-Systeme vor Ort zwingend gefordert. Diese Systeme basieren grundlegend auf den Konzepten gemäß der CEUS-Referenzarchitektur. Jede Universität hat jedoch entsprechend ihrer individuellen Anforderungen und Spezifika unterschiedliche Schwerpunkte bei der Ausgestaltung und Weiterentwicklung ihres lokalen DWH-Systems gesetzt. Obwohl für die oben genannten Domänen eine landesweit weitgehend homogene operative IT-Landschaft vorliegt, entstanden nennenswerte Unterschiede in den darauf aufbauenden DWH-Systemen. Um den individuellen Informationsbedarf einer Universität bedienen zu können, wurden fortwährend die jeweiligen Schnittstellen zu deren operativen Systemen sowie die daran angeschlossenen ETL-Prozesse (Extract, Transform, Load; vgl. z. B. [BG04, S. 49 ff.]) des lokalen DWH-Systems modifiziert. Dementsprechend wurden auch die zugehörigen Datenbank- und Analyse-Schemata an diese Änderungen oder Erweiterungen angepasst. Der isolierte Betrieb der lokalen DWH-Systeme erforderte zudem die Bereitstellung entsprechend qualifizierter Personen an den einzelnen Universitäten vor Ort. Synergieeffekte konnten daher nur zum Teil genutzt und Ressourcenersparnisse bei Pflege und Wartung der Systeme lediglich in geringem Umfang erzielt werden. Die eigenständige Verwaltung der lokalen DWH-Systeme weist zudem Nachteile insbesondere bei der Implementierung allgemeingültiger Anforderungen im Berichtswesen auf (z. B. landesweite Berichtsanforderungen vom StMWFK). Jedes DWH-System muss hierzu für sich angepasst und erweitert werden. Der parallele Betrieb von isolierten DWH-Systemen für einen einzigen Aufgabenbereich führte im Projektverlauf letztlich zu ähnlichen Aufwänden, wie der Betrieb von DWH-Systemen unterschiedlicher Aufgabenbereiche.

In einer weiteren Ausbaustufe werden seit 2006 die bayerischen Fachhochschulen in die landesweite CEUS-Gesamtarchitektur mit eingebunden. Den einzelnen Fachhochschulen soll hierbei jeweils ein lokales DWH-System, zunächst mit Detaildaten zu Studierenden und Prüfungen, zur Verfügung gestellt werden. Dabei sollen Synergieeffekte bei der Konzeption sowie beim Betrieb und der Weiterentwicklung dieser DWH-Systeme bestmöglich genutzt werden. Die bei den bayerischen Universitäten verfolgte Strategie, eigenständige DWH-Systeme anhand eines Referenzmodells zu realisieren, erschien bei den Fachhochschulen aus wirtschaftlichen Gründen (höhere Anzahl, durchschnittliche Größe der Fachhochschulen etc.) und den genannten Erfahrungen als nicht adäquat.

Aufgrund der oben geschilderten Rahmenbedingungen ergeben sich die nachfolgend dargestellten Anforderungen des IHB an die lokalen DWH-Systeme der Fachhochschulen:

- *Standardisierte Schnittstelle:* Die Datenextraktion soll durch die Fachhochschulen gemäß einer einheitlichen, verbindlichen Schnittstelle erfolgen. Hierdurch soll Unabhängigkeit von den jeweiligen operativen Systemen erreicht werden. Auf diese Weise wird ebenfalls gewährleistet, dass personenbezogene Daten im Vertrauensbereich der Fachhochschulen extrahiert und vor der Übergabe an die DWH-Systeme eliminiert oder anonymisiert werden. Dies erlaubt die Einhaltung datenschutzrechtlicher Vorgaben. In einem DWH-System ist der

Datenschutzproblematik besondere Aufmerksamkeit zu widmen, da die dort angebotenen Navigationsoperatoren auf einer umfassenden Datenbasis vielfältige Kombinationsmöglichkeiten von Einzeldaten ermöglichen (vgl. z. B. [Mö99], [SPU02]).

- *Einheitliches Datenbankschema*: Als Voraussetzung für einen einheitlichen ETL-Prozess soll ein einheitliches Datenbankschema konzipiert werden, das die Informationsbedarfe aller Fachhochschulen erfüllt.
- *Einheitlicher ETL-Prozess*: Die zu entwickelnden lokalen DWH-Systeme sollen über einen einheitlichen ETL-Prozess versorgt werden, da der Betrieb und die Wartung von 17 eigenständigen Ladeprozessen im Rahmen der verfügbaren personellen Ressourcen nicht effizient durchführbar sind. Auch bei einer sukzessiven Einführung der einzelnen Systeme kann auf diese Weise ein homogener ETL-Prozess nachhaltig gewährleistet werden. Anpassungen und Weiterentwicklungen sind entsprechend nur einmalig an einer Stelle durchzuführen.
- *Zentrale Steuerung des ETL-Prozesses*: Der einheitliche ETL-Prozess soll von zentraler Stelle administriert und gesteuert werden können. Die Steuerung des Prozesses ist durch eine Parametrisierung zu ergänzen, um die DWH-Systeme einzeln, domänenspezifisch und zeitspezifisch laden zu können.
- *Standardisiertes Berechtigungskonzept*: Das Berechtigungskonzept umfasst eine genaue Definition der Zugriffsberechtigungen auf die Datenbasis des DWH-Systems. Um dessen Pflege zu erleichtern, sollen die Berechtigungen mittels eines für die einzelnen Fachhochschulen gleichartig definierten Rollenkonzepts verwaltet werden.

In Zusammenarbeit mit den Fachhochschulen kommen darüber hinaus folgende Anforderungen an die Konzeption des DWH-Systems hinzu:

- *Hochschulübergreifendes Berichtswesen*: Um die Vergleichbarkeit des Berichtswesens im Hinblick auf die Zielvereinbarungen mit dem StMWFK sicherzustellen, soll ein fachhochschulübergreifendes Standardberichtswesen geschaffen werden. Dies erfordert, die Semantik der Kennzahlen und Merkmale für alle Fachhochschulen eindeutig festzulegen.
- *Hochschulspezifisches Berichtswesen*: Neben dem hochschulübergreifenden Standardberichtswesen sollen auch spezifische Anforderungen berücksichtigt werden können, um den individuellen Informationsbedarf der Fachhochschulen zu decken.
- *Minimierung der Implementierungs- und Betriebskosten*: Der Implementierungs-, Wartungs- und Pflegeaufwand ist auf Seiten der Fachhochschulen möglichst gering zu halten. Der Betrieb der DWH-Systeme soll daher nicht an jeder Fachhochschule separat erfolgen, da hierfür redundant Hardware, Software und Personalressourcen vorgehalten werden müssen.

Im folgenden Kapitel wird – auf Grundlage der genannten Anforderungen – die vom ihm entwickelte Architektur für die lokalen DWH-Systeme der bayerischen Fachhochschulen vorgestellt.

3 Konzeption eines mandantenfähigen DWH-Systems

Mit Bezug zu den vorausgegangenen Ausführungen wird nachfolgend die Architektur des entwickelten DWH-Systems vorgestellt. Aufgrund der in Kapitel 2 genannten Anforderungen stellten sich eigenständige DWH-Systeme für die Fachhochschulen als ungeeignet heraus. Aus diesem Grund wurde nach Konzepten für eine zentrale Architektur gesucht. Das in der Praxis bei transaktionalen Systemen bewährte Prinzip der Mandantenfähigkeit lässt sich auf die Konzeption eines DWH-Systems für die Fachhochschulen übertragen. Im Folgenden wird zunächst die Grundidee eines Mandantensystems eingeführt und dessen Eignung für die Umsetzung der hier vorliegenden, spezifischen Anforderungen aufgezeigt. Anschließend wird der Begriff eines mandantenfähigen DWH-Systems definiert und die vom ihm entwickelte mandantenfähige DWH-Architektur vorgestellt.

Nach Schulze wird als Mandantensystem Software bezeichnet, die so konzipiert ist, dass gleichartige Arbeiten für unterschiedliche Mandanten durchgeführt werden können [Sc89, S. 1843]. Im SAP-Umfeld kommt der Begriff des Mandanten dem Begriff eines Konzerns nahe. Der Mandant im System R/3 bildet eine logische Gruppierung von handelsrechtlichen, organisatorischen und administrativen Einheiten und ist damit die höchste Organisationseinheit [EGN96, S. 45]. Auf diese Weise lassen sich in einem gemeinsamen System mehrere eigenständige Unternehmen getrennt voneinander abbilden. Die datentechnische Abtrennung wird hierbei durch die Aufnahme des Mandanten als Schlüssel in die Tabellen des R/3-Repositorys erreicht [Wi98, S. 26]. Im Allgemeinen kann in einem Mandantensystem zwischen mandantenabhängigen und mandantenunabhängigen Strukturen unterschieden werden. Mandantenabhängige Strukturen werden für jeden Mandanten individuell gestaltet. Mandantenunabhängige Strukturen sind übergreifend für alle am System Beteiligten verfügbar.

Der Begriff des Mandanten und die in diesem Zusammenhang beschriebene datentechnische Abtrennung lassen sich auf die Fachhochschulen übertragen, die weitgehend eigenständig und autonom agieren. Die Grundidee eines Mandantensystems wird im Folgenden für die Konzeption der lokalen DWH-Systeme aufgegriffen und auf die Spezifika eines DWH-Systems im Hochschulumfeld ausgerichtet.

Im Rahmen dieser Arbeit wird Mandantenfähigkeit eines DWH-Systems folgendermaßen definiert:

Ein mandantenfähiges Data-Warehouse-System ist ein Data-Warehouse-System, das für eigenständige Organisationseinheiten entscheidungsrelevante Daten aus unterschiedlichen Datenquellen über einen einheitlichen ETL-Prozess in einer konsolidierten Datenbasis bereitstellt und das darauf aufbauende multidimensionale Schema sowohl mandantenabhängige als auch mandantenunabhängige Strukturen enthält.

An dieser Stelle ist darauf hinzuweisen, dass zum einen die Quellsysteme der beteiligten Mandanten einem logischen Aufgabenbereich entstammen sollten. Die Integration stark heterogener Quellsysteme im Sinne der abgebildeten Diskurswelt führt zu einer hohen

Anzahl mandantenabhängiger Strukturen und würde die Vorteile eines mandantenfähigen DWH-Systems egalisieren. Zum anderen ist für dessen Einführung eine hohe Übereinstimmung bzgl. der Analyseanforderungen der einzelnen Mandanten (z. B. ein weitgehend homogenes Merkmals- und Kennzahlensystem) eine entscheidende Voraussetzung.

Die logische Architektur eines DWH-Systems lässt sich in die Ebenen *Datenerfassung*, *Datenhaltung* und *Datenbereitstellung* gliedern (vgl. z. B. [BU00, S. 17 ff.]). Die Datenerfassungsebene stellt Werkzeuge zur Extraktion der Daten aus den operativen Systemen sowie zur Datenbereinigung zur Verfügung. Das eigentliche Data-Warehouse auf der Datenhaltungsebene bildet eine konsolidierte Datenbasis. Die Datenbereitstellungsebene besteht aus einem OLAP-Server, der die multidimensional aufbereiteten Daten den Präsentationswerkzeugen zur Verfügung stellt (vgl. z. B. [BU00, S. 2 ff.]). Im Folgenden wird das vom ihb entwickelte mandantenfähige DWH-System entlang dieser Ebenen erläutert. Hierbei werden insbesondere die Anpassungen bzw. Erweiterungen im Vergleich zu einem nicht mandantenfähigen Konzept aufgezeigt.

3.1 Datenerfassungsebene

Auf der Datenerfassungsebene ist eine einheitliche Schnittstellenbeschreibung für die Anbindung der operativen Systeme zu erarbeiten. Diese bildet den Ausgangspunkt für einen gemeinsamen Ladeprozess aller Mandanten. Die zu definierende Schnittstelle muss zum einen die informationellen Anforderungen aller beteiligten Mandanten abdecken und zum anderen die Bedeutung der einzelnen Elemente der Schnittstelle eindeutig bestimmen. Durch sie werden die Extraktion sowie eine Teilaufgabe der Homogenisierung der Daten an den Datenlieferanten ausgelagert. Die Schnittstellendefinition ist für alle Mandanten verpflichtend. Sie umfasst mandantenübergreifende Elemente, die von allen Mandanten gefüllt werden, sowie Felder, die individuelle Erweiterungen darstellen und nicht von allen Mandanten bedient werden müssen.

Die Datenextraktion findet im Vertrauensbereich der einzelnen Mandanten statt. Diese bestimmen selbstständig den Zeitpunkt sowie die Häufigkeit der Extraktion aus den eigenen operativen Systemen und sind für die Qualität der Daten verantwortlich. Neben der Datenextraktion erfolgt eine Datentransformation entsprechend der Konventionen der Schnittstellenbeschreibung. Hierbei werden u. a. auch mandantenspezifische Identifikationsnummern hinzugefügt und alle personenbezogenen Angaben eliminiert oder anonymisiert. Die verschlüsselten Daten werden im Anschluss über eine gesicherte Verbindung per Push-Prinzip an einen zentralen Upload-Server (vgl. Abb. 1) übertragen.

Über einen zentralen ETL-Prozess werden die von den einzelnen Mandanten gemäß der Schnittstellenbeschreibung bereitgestellten Daten in das Data-Warehouse geladen. Durch die Vorgabe einer einheitlichen und verbindlichen Schnittstelle wird der Ladeprozess erheblich vereinfacht, da ein Teil der syntaktischen und semantischen Homogenisierung der Daten bereits im Verantwortungsbereich des Mandanten erfolgt. Die notwendigen Schritte zur Datentransformation sind zudem nur einmal zu spezifizieren und gelten für alle Mandanten gleichermaßen. Der ETL-Prozess ist darüber hinaus parametrisierbar, um abhängig von Mandant, Domäne und Zeitpunkt in unterschiedlichen Varianten aus-

geführt werden zu können. Für die Administration steht ein eigener Steuerungs- und Überwachungsprozess zur Verfügung. Dieser regelt die jeweilige Kombination der oben genannten Parameter.

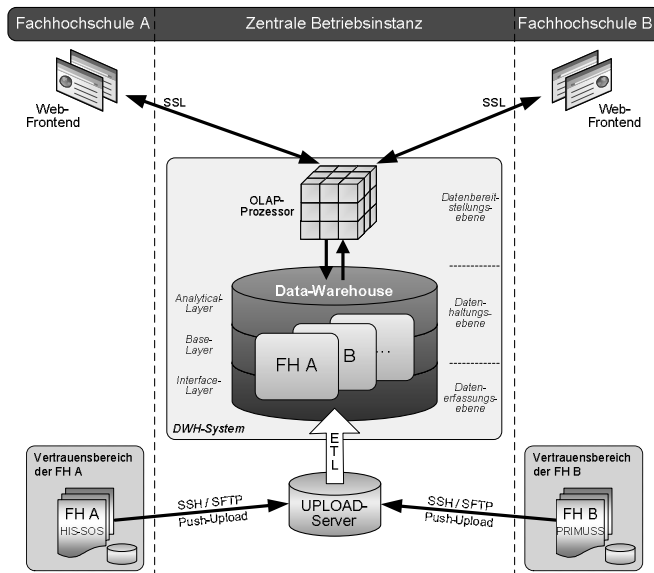


Abbildung 1: Architektur des mandantenfähigen DWH-Systems für die bayerischen Fachhochschulen

3.2 Datenhaltungsebene

Die Datenhaltungsebene des mandantenfähigen DWH-Systems ist in drei Ebenen gegliedert (vgl. z. B. [Si01, S. 12 ff.]). Der Interface-Layer ist als temporärer Speicher konzipiert, dessen Datenstruktur bei der vorliegenden Architektur exakt der Schnittstellendefinition entspricht. Eine konsistente, normalisierte Datenhaltungsschicht bildet der Base-Layer, der die konsolidierten und historisierten Daten permanent speichert. Der Analytical-Layer ist für die Bereitstellung der Daten in der von den jeweiligen Anfrage-Werkzeugen benötigten Form vorgesehen. Im Gegensatz zur Konzeption mehrerer eigenständiger DWH-Systeme sind bei einem mandantenfähigen Ansatz diese Ebenen mit jeweils einem mandantenübergreifenden Datenschema nur einmal vorhanden.

Ein Datenschema besteht hier grundsätzlich aus mandantenabhängigen und mandanten-unabhängigen Schemaobjekten. Beispielsweise ist das *Semester* fachhochschulübergreifend definiert und damit mandantenunabhängig. Beim *Studiengang* hingegen handelt es sich um ein mandantenabhängiges Objekt, da an jeder Fachhochschule andere Studiengänge angeboten werden. Unabhängig davon, ob eine einzelne Fachhochschule ein bestimmtes Schemaobjekt verwendet oder nicht, steht jedes allen Mandanten gleichermaßen zur Verfügung. So wird der *Schwerpunkt* nur von einem Teil der Fachhochschulen für Auswertungen benötigt und daher auch nur von diesen Hochschulen mit Daten

versorgt. Um die Daten später mandantenspezifisch auswerten zu können, wird der Mandant als Teil des Schlüssels in alle mandantenabhängigen Schemaobjekte aufgenommen. Über diese Schemaobjekte können somit bis zu einem gewissen Grad individuelle Anforderungen abgebildet werden. Des Weiteren stehen durch das gemeinsame Datenschema Erweiterungen in der Implementierung allen am System Beteiligten gleichzeitig zur Verfügung, wodurch sich die Pflege und Wartung des Systems vereinfacht.

3.3 Datenbereitstellungsebene

Die Datenbereitstellungsebene bildet die Schnittstelle zu den Endanwendern. Ein OLAP-Server stellt die multidimensionalen Strukturen zur Verfügung und ermöglicht den interaktiven Zugriff auf das Data-Warehouse. Der Zugriff auf das Analyseschema wird für alle Mandanten über ein einheitliches Berechtigungskonzept realisiert. Dazu sieht das Analyseschema eine Dimension zur Aufnahme des Mandanten vor. Diese Dimension wird im Berechtigungskonzept genutzt um die datentechnische Abtrennung der einzelnen Mandanten durch entsprechende Filterkriterien zu garantieren. Hierdurch sind mandantenspezifische Inhalte nur für den jeweiligen Mandanten verfügbar.

Das Analyseschema beruht auf einem multidimensionalen Datenschema. Dieses umfasst sowohl mandantenabhängige als auch mandantenunabhängige Dimensionen und Kennzahlen. Generell stehen die Dimensionen und Kennzahlen für alle Mandanten gleichermaßen zur Verfügung. Per Berechtigungskonzept kann einem Mandanten jeweils nur auf die für ihn relevanten Dimensionen und Kennzahlen Zugriff gegeben werden. Insgesamt ergibt sich die Notwendigkeit in Bezug auf Benennung, Definition und Abhängigkeiten der Schemaobjekte einen Konsens zwischen den Mandanten zu erarbeiten. Dies genügt dem Anspruch, die Vergleichbarkeit der Schemaobjekte sicherzustellen. Zudem bildet das gemeinsame Analyseschema die Grundlage für den Aufbau eines standardisierten Berichtswesens.

4 Realisierung des mandantenfähigen DWH-Systems

Im Jahr 2006 begann das ibh mit der Entwicklung des mandantenfähigen Prototyps für die Domäne *Studierenden und Prüfungen*. Wichtigstes Ziel des ibh war dabei, den Implementierungs- und Wartungsaufwand für die sukzessive Einführung an den 17 bayerischen Fachhochschulen so gering wie möglich zu halten. Das umgesetzte mandantenfähige Konzept erforderte initial einen relativ hohen Entwicklungsaufwand. Der anschließende Aufwand für die Realisierung der einzelnen Mandanten ist jedoch im Vergleich zu einem nicht-mandantenfähigen Konzept deutlich geringer (vgl. Abb. 2).

Basierend auf dem im vorangegangenen Kapitel erarbeiteten Konzept eines mandantenfähigen DWH-Systems wird nachfolgend dessen Realisierung für die lokalen DWH-Systeme der bayerischen Fachhochschulen vorgestellt. Die zentrale Betriebsinstanz (vgl. Abb. 1) wird derzeit durch das ibh bereitgestellt. Diese besteht aus einer gemeinsamen Hardware-Architektur für alle Fachhochschulen. Die Datenlieferungen erfolgen über

einen zentralen Upload-Server. Für die gesicherte Übertragung wird hierbei das SSH-Protokoll genutzt. Als Datenbankverwaltungssystem kommt der Microsoft SQL Server 2005 zum Einsatz. Bei der Implementierung des ETL-Prozesses wird auf die SQL Server Integration Services (SSIS) zurückgegriffen. Die Datenbereitstellung wird mit Hilfe des MicroStrategy Intelligence Servers realisiert. Der Zugriff auf das Berichtswesen findet dabei ausschließlich über das MicroStrategy Web-Frontend (SSL-verschlüsselt) statt.

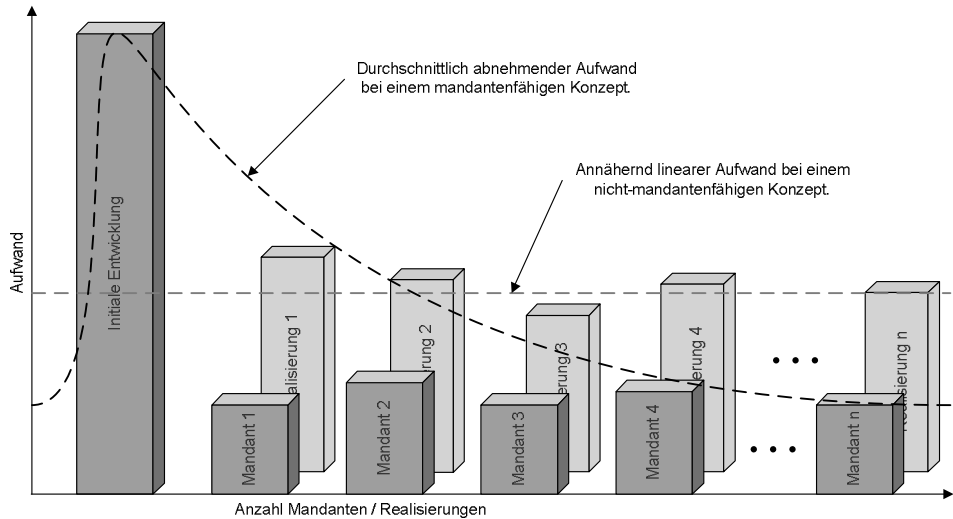


Abbildung 2: Gegenüberstellung der Realisierungsaufwände für ein DWH-System nach einem mandantenfähigen und einem nicht-mandantenfähigen Konzept

Die nachstehenden Ausführungen greifen die Unterteilung in Datenerfassungs-, Datenhaltungs- und Datenbereitstellungsebene wieder auf.

4.1 Datenerfassungsebene

Die Verwaltung der Studierenden- und Prüfungsdaten erfolgt bei den bayerischen Fachhochschulen entweder mithilfe des Anwendungsprogramms HIS-SOS/POS oder PRIMUSS. Die den beiden Systemen zugrunde liegenden Datenstrukturen unterscheiden sich sowohl auf syntaktischer als auch semantischer Ebene. Um die konzeptionell vorgesehene einheitliche Schnittstelle zu ermöglichen und einen gemeinsamen ETL-Prozess realisieren zu können, wurden auf der Datenerfassungsebene folgende Vorbereitungen getroffen:

- Die Heterogenität der Datenstrukturen von HIS-SOS/POS und PRIMUSS wurde bei der Analyse der jeweiligen Datenschemata hinsichtlich Gemeinsamkeiten und Differenzen aufgelöst. Basierend auf diesen Ergebnissen wurde die für beide Vorsysteme einheitliche und verbindliche Schnittstellenbeschreibung festgelegt. Diese enthält neben allgemeinen Anweisungen zur Datenextraktion und -lieferung für jede Satzart die exakte Struktur und das genaue Format, in dem die Daten bereitzustellen sind. In jeder Satzart ist zudem ein Feld vorgesehen, um die jeweilige Fachhochschule (über die Hochschulnummer) zu identifizieren.
- Die Einhaltung datenschutzrechtlicher Vorgaben kann aufgrund der detaillierten Schnittstellenbeschreibung durch den örtlichen Datenschutzbeauftragten gut überprüft werden. Felder, die direkt personenbezogene Daten enthalten (z. B. Name, Adresse, Telefonnummer), sind in der für die Fachhochschulen entwickelten Satzartenbeschreibung nicht enthalten. Identifizierende Merkmale (z. B. Matrikelnummer), die zur korrekten Verknüpfung der Daten aus unterschiedlichen Satzarten erforderlich sind, werden u. a. mittels eines vom ihb bereitgestellten Werkzeugs verschlüsselt (z. B. RSA-Algorithmus).
- Jede Fachhochschule stellt ihre Studierenden- und Prüfungsdaten gemäß der Satzartenbeschreibung als so genannte Flatfiles auf dem zentralen Upload-Server zur Verfügung (Push-Prinzip). Auf diese Weise bestimmen die Fachhochschulen zugleich die Aktualisierungsstrategie für ihre Daten im Data-Warehouse.
- Bei der Ausgestaltung des für alle Fachhochschulen im Wesentlichen einheitlichen ETL-Prozesses galt es, neben den allgemeingültigen Übertragungsregeln auch einige wenige logische und semantische Besonderheiten einzelner Mandanten zu berücksichtigen. Bei mandantenspezifischen Daten werden daher bei jedem Transformationsschritt der jeweilige Mandant und dessen Vorsystem als Parameter mit eingebunden. Auf diese Weise können Spezifika (z. B. unterschiedliche Zuordnung von gemeinsamen Schwerpunktfächern zu Studiengängen) explizit behandelt und auf das gemeinsame Datenschema abgebildet werden.
- Um den über SSIS-Pakete abgebildeten ETL-Prozess parametrisieren zu können, wurden Tabellen zur Steuerung des gesamten Ladeprozesses erstellt. Diese beinhalten Angaben zu Fachhochschule, Domäne (z. B. Studierendendaten) sowie Zeitraum der zu ladenden Daten. Die Protokollierung des Ladeprozesses wurde dabei so realisiert, dass anhand eines Identifiers in der Steuerungstabelle die einzelnen Ladeschritte sowie ggf. aufgetretene (Daten-)Fehler automatisiert aufgelistet und über Fehlerlisten per E-Mail an die Administratoren versendet werden.
- Die Transformations- und Ladeprozesse für Interface-, Base- und Analytical-Layer können individuell angesteuert werden. Ein übergeordnetes SSIS-Paket erlaubt den durchgängigen Ablauf des gesamten ETL-Prozesses.

4.2 Datenhaltungsebene

Zur Realisierung der Datenhaltungsebene wurde für die Interface-, Base- und Analytical-Layer jeweils ein Schema angelegt. Diese dienen der Speicherung der Daten aller Fachhochschulen. Im Folgenden werden die zur Realisierung der Mandantenfähigkeit notwendigen Anpassungen aufgezeigt (vgl. Abb. 3):

- Das Schema des Interface-Layer entspricht genau den Strukturen der Satzarten. Die von den Fachhochschulen bereitgestellten Daten werden durch die SSIS-Pakete ohne Transformation in den Interface-Layer übertragen. Dabei stellt eine Tabelle einen temporären Speicher für die Daten einer Satzart aller beteiligten Fachhochschulen dar. Das Kennzeichen Hochschulnummer dient der Unterscheidung der mandantenspezifischen Daten.
- In allen Tabellen des Base-Layer mit mandantenspezifischen Daten wurde die Hochschulnummer als Teil des identifizierenden Schlüssels aufgenommen. Dadurch können die Daten einer Hochschule einzeln und unabhängig von anderen Hochschulen verwaltet werden. In Kombination mit den in allen Schlüsseltabellen vorgesehenen Zeitstempeln, kann darüber hinaus eine individuelle Historisierung der Datensätze realisiert werden. Zur Durchführung der korrekten, mandantenspezifischen Historisierung sowie der Datentransformation wurden für den Übergang von Interface- auf Base-Layer entsprechende Sichten angelegt, welche die vollständige Historisierungs- bzw. Transformationslogik abbilden.
- Analog zum Base-Layer führen auch im Analytical-Layer alle Fakttabellen und mandantenabhängigen Dimensionstabellen das Attribut Mandant mit sich. Dies bildet einen elementaren Baustein für das Berechtigungskonzept, um einer Fachhochschule exklusiven Zugriff auf ihre Stamm- und Bewegungsdaten zu gewährleisten.
- Zur Verarbeitung der Daten durch den auf dem ROLAP-Prinzip (vgl. z. B. [BU00, S. 27 f.]) basierenden Intelligence Server von MicroStrategy wurde dem Analytical-Layer ein denormalisiertes Snowflake-Schema zugrunde gelegt.

Aufgrund der Implementierung eines zentralen, mandantenfähigen DWH-Systems, können die Aufgaben zur Administration des Datenbankverwaltungssystems für alle Fachhochschulen einheitlich und zeitgleich erfüllt werden (z. B. Verwaltung der Indizes, Datenbanksicherungen, Archivierung der Datenbestände).

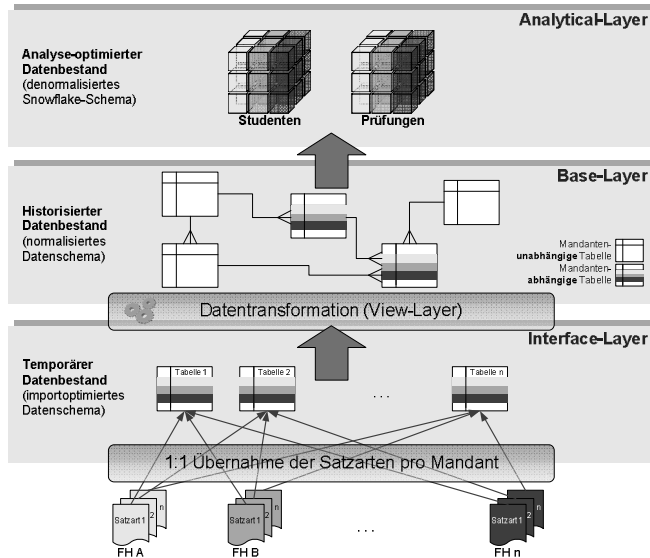


Abbildung 3: Logische Architektur des mandantenfähigen Data-Warehouse

4.3 Datenbereitstellungsebene

Der auf der Datenbereitstellungsebene durch den MicroStrategy Intelligence Server unterstützte Analyseprozess operiert auf den Daten des Analytical-Layer. Die Daten werden den Endanwendern über fachbezogene Schemaobjekte präsentiert. Die Komplexität der Datenstrukturen eines Data-Warehouse bleibt daher für den Endanwender hinter der semantischen Schicht der Terminologie verborgen [LJ99, S. 7-1 f.]. Die Besonderheiten der mandantenspezifischen Implementierung der Datenbereitstellungsebene werden nachstehend erörtert:

- Analog zur Datenhaltungsebene musste ein mandantenspezifischer Zugriff auf der Datenbereitstellungsebene realisiert werden. Im Analyseschema wurde daher ein Schemaobjekt zur Aufnahme des Mandanten angelegt. Jeder hochschulspezifischen Anwendergruppe wird dieses Schemaobjekt mit der jeweils entsprechenden Ausprägung (hier die Hochschulnummer) als Sicherheitsfilter hinterlegt. Anwender sind jeweils nur einer Anwendergruppe zugeordnet. Auf diese Weise wird gewährleistet, dass nur die für eine Fachhochschule relevanten Schemaobjekte zur Datenanalyse bereitgestellt und nur die mandanteneigenen Daten in einem Bericht ausgegeben werden.
- Um das Berechtigungskonzept leicht administrieren zu können, wurden fachhochschulübergreifend gültige Zugriffsrollen (Standard-, Advanced-, und Power-User) entworfen. Diese unterscheiden sich hinsichtlich der auf den Datenbeständen erlaubten Operationen (von der einfachen Berichtsausführung bis hin zur eigenständigen Berichtsdefinition). Jedem Endanwender wird entsprechend seiner Profildimension (vgl. [HM07, S. 582]) eine Rolle zugewiesen.

- Aufbauend auf der für alle Fachhochschulen einheitlichen Definition der Dimensionen und bestimmter Kennzahlen, erfolgte die Ausgestaltung eines konsolidierten, übergreifenden Standardberichtswesens. Der Aufbau und die Semantik der von allen Fachhochschulen benötigten Berichte wurden in einer Arbeitsgruppe, bestehend aus Fachverantwortlichen der einzelnen Fachhochschulen, gemeinsam definiert und anschließend vom ihb implementiert. Zur Ausführungszeit eines Berichts werden über den Sicherheitsfilter des Endanwenders die jeweils zugehörigen Dateninhalte ermittelt. Die identischen Berichtsstrukturen erlauben auf diese Weise eine Vergleichbarkeit der fachhochschulspezifischen Berichtsergebnisse.
- Um die individuellen Informationsbedarfe vollständig zu erfüllen, kann jede Fachhochschule zudem eigene Berichte erstellen. Die fachlichen Gegebenheiten einer Fachhochschule werden durch die mandantenspezifischen Anteile in der Behandlung der Daten auf Datenerfassungs-, Datenhaltungs- und Datenbereitstellungsebene abgebildet. Neben einem standardisierten Berichtswesen steht damit jeder Fachhochschule eine flexible Informationsversorgung zum Aufbau eines personalisierten Informationsportals zur Verfügung.

Die gewonnenen Erkenntnisse im Rahmen der Realisierung des mandantenfähigen DWH-Systems für die Fachhochschulen legen offen, dass ein entscheidendes Kriterium bei der Ausgestaltung eines mandantenfähigen DWH-Systems in der Differenzierung zwischen mandantenabhängigen und mandantenunabhängigen Schemaobjekten liegt. Trotz des bei den Fachhochschulen vorliegenden gleichen Domänenbezugs entstand während der Konzeptionsphase ein nicht unerheblicher Aufwand bei der Bewertung und Definition der Stammdaten als mandantenabhängig bzw. mandantenunabhängig.

Mit steigender Anzahl von Schemaobjekten, die nur von einem Teil der Mandanten verwendet werden, sinkt der Mehrwert, der mit einem derartigen DWH-Konzept erzielbar ist. Davon abhängig ist zu entscheiden, ob die Implementierung eines mandantenfähigen DWH-Systems unter betriebswirtschaftlichen Gesichtspunkten sinnvoll ist. Grundsätzlich besteht hier ein Trade-off zwischen der Berücksichtigung individueller Anforderungen der Mandanten und der insgesamt potentiell realisierbaren Betriebskostensparnis.

5 Zusammenfassung

Der vorliegende Beitrag erläutert die Anforderungen an die Architektur der lokalen DWH-Systeme für die bayerischen Fachhochschulen. Als Lösung wurde ein mandantenfähiges DWH-System mit speziellen Architekturmerkmalen vorgestellt, das am ihb konzipiert und entwickelt wurde. Zu den wesentlichen Aufgaben des ihb bei der Einführung der lokalen DWH-Systeme zählten die Unterstützung der Fachhochschulen bei der Bereitstellung der Daten gemäß der Satzartenbeschreibung, die Prüfung der gelieferten Daten auf Konsistenz und Plausibilität, das Laden der Daten sowie die Rückmeldung über aufgetretene Datenfehler, die Anpassung und Weiterentwicklung des MicroStrategy-Schemas und die Unterstützung bei der Berichtsdefinition. Das mandantenfähige

DWH-System befindet sich gegenwärtig für fünf Fachhochschulen im Betrieb. Die sukzessive Einbindung weiterer Fachhochschulen erfolgt parallel.

Die Mandantenfähigkeit des DWH-Systems wirkt sich insbesondere vorteilhaft durch den reduzierten Aufwand für Pflege und Wartung sowie die gemeinsame Nutzung von technischen und personellen Ressourcen für alle Fachhochschulen aus. Die Architektur ermöglicht zudem eine leichte Einbindung weiterer Fachhochschulen. Dies wird insbesondere durch den gemeinsamen Ladevorgang (ETL-Prozess) und das gemeinsame Datenbankschema erreicht. Nachteilig zeigte sich in der Konzeptionsphase vor allem der notwendige Abstimmungsaufwand zwischen den Fachhochschulen. Darüber hinaus sind zukünftige Änderungen zwischen allen Mandanten zu koordinieren.

Die bayerischen Fachhochschulen erhalten durch das vorgestellte mandantenfähige DWH-System ein zukunftssicheres und erweiterbares Anwendungssystem, mit dem auf spezifische Informationsbedarfe individuell reagiert werden kann. Es steht eine standardisierte, flexible Informationsversorgung durch ein umfangreiches Berichtswesen (hochschulübergreifend und hochschulspezifisch) zur Verfügung. Dies ermöglicht eine gezielte und schnelle Informationsrecherche bei einheitlicher Begriffsdefinition sowie eine Vergleichbarkeit der Berichte. Hierdurch konnte bei den Fachhochschulen in wenigen Monaten eine hohe Akzeptanz des Systems CEUS erzielt werden. Dies zeigt sich auch anhand folgender Punkte: Die zunächst prototypische Implementierung konnte erfolgreich in den Systembetrieb überführt werden. Des Weiteren wurde das System innerhalb von zwei Monaten drei weiteren Fachhochschulen zur Verfügung gestellt. Auch der Nutzerkreis innerhalb einer Fachhochschule wird permanent erweitert (z. B. Roll-out an den Fakultäten). Zudem haben bereits mehrere Fachhochschulen ein starkes Interesse an einer Einbindung in die mandantenfähige DWH-Lösung von CEUS bekundet.

6 Literaturverzeichnis

- [BG04] Bauer, A.; Günzel, H.: Data-Warehouse-Systeme – Architektur, Entwicklung, Anwendung. dpunkt Verlag, Heidelberg, 2004.
- [BU00] Böhnlein, M.; Ulbrich-vom Ende, A.: Grundlagen des Data Warehousing: Modellierung und Architektur. Bamberger Beiträge zur Wirtschaftsinformatik Nr. 55, Bamberg, 2000.
- [EGN96] Engels, A.; Gresch, J.; Nottenkämper, N.: SAP R/3 kompakt – Einführung und Arbeitsbuch für die Praxis. Tewi Verlag, München, 1996.
- [HM07] Henrich, A.; Morgenroth, K.: Kontextbasiertes Information Retrieval zur Unterstützung von Entwicklungsprozessen. In (Oberweis, A.; Weinhardt, C. et al. Hrsg.): eOrganisation – Service-, Prozess-, Market-Engineering: 8. Internationale Tagung Wirtschaftsinformatik – Band 2. Universitätsverlag, Karlsruhe, 2007; S.575-592.
- [Ku08] Kurz, D.: New Public Management und Verwaltungskomplexität. In (von Maravic, P.; Priddat, B.P. Hrsg.): Öffentlich – Privat: Verwaltung als Schnittstellenmanagement. Metropolis Verlag, Marburg, 2008; S. 211-232.

- [LJ99] Lehmann, P.; Jaszewski, J.: Business Terms as a Critical Success Factor for Data Warehousing. In: Proceedings of the International Workshop on Design and Management of Data Warehouses (DMDW'99), Heidelberg, 1999; S.7-1 – 7-5.
- [Mö99] Möncke, U.: Sicherheit im Data Warehouse – Profilbildung und Anonymität. In (Hors-ter, P.; Fox, D. Hrsg.): Datenschutz und Datensicherheit – Konzepte, Realisierungen, Rechtliche Aspekte, Anwendungen. Vieweg, Braunschweig/Wiesbaden 1999, S. 30-59.
- [Sc89] Schulze, H.H.: Computer Enzyklopädie. Rowohlt Taschenbuch Verlag, Hamburg, 1989.
- [Si01] Sinz, E.J. et. al.: Architekturkonzept eines verteilten Data-Warehouse-Systems für das Hochschulwesen. Bamberger Beiträge zur Wirtschaftsinformatik Nr. 59, Bamberg, 2001.
- [SPU02] Sinz E.J., Plaha M., Ulbrich-vom Ende A.: Datenschutz und Datensicherheit in einem landesweiten Data-Warehouse-System für das Hochschulwesen. Bamberger Beiträge zur Wirtschaftsinformatik und Angewandten Informatik Nr. 62, Bamberg, 2002.
- [Wi98] Will, L.: SAP R/3-Systemadministration: Basiswissen für das SAP R/3-Systemma-nagement. Addison-Wesley-Longman, Bonn, 1998.

Entwicklung eines Management Cockpits mit SAP BI 7.0: ein Erfahrungsbericht

Jörg Raimann*, Simone Schnee**, Nina Gudelj***, Thomas Broda*

*SAP IS/U, CRM und Business Warehouse
Mainova AG
Solmsstr. 38
60623 Frankfurt am Main
{j.raimann | t.broda}@mainova.de

**Controlling und Unternehmenssteuerung
Mainova AG
s.schnee@mainova.de

***Fachhochschule Frankfurt am Main
Nibelungenplatz 1
60318 Frankfurt am Main

Abstract: Der Praxisbeitrag beschreibt die Konzeption und Realisierung einer Management Cockpit-Anwendung beim Energieversorgungsunternehmen Mainova AG, das zentrale Kennzahlen auf Basis der für das Unternehmen maßgeblichen Steuerungsebenen in integrierter und visuell aufbereiteter Form darstellt. Relevant für die Unternehmenssteuerung ist zum einen die Steuerungsebene nach Wertschöpfungsstufen; zum anderen wird diese Ebene um eine Segmentsicht ergänzt. In der Einführung wird zunächst das dem Beitrag zugrunde liegende Verständnis des Begriffs „Management Cockpit“ erläutert. Nach einem Überblick über die Projekthistorie sowie die wesentlichen Ziele werden im darauf folgenden Abschnitt die Anforderungen an das Projekt aus betriebswirtschaftlicher Sicht dargestellt. Im Einzelnen werden das dem Management Cockpit zugrunde liegende Steuerungskonzept, die Auswahl und Systematik der abgebildeten Kennzahlen sowie die Investitions- und Kostenprojektplanung als ein wesentlicher Baustein für das Management Cockpit näher beschrieben. Im Rahmen des Abschnitts IT-Umsetzung wird zunächst auf das allgemeine Vorgehen im Projekt eingegangen. Die IT-Architektur, basierend auf den mit SAP Netweaver 7.0 verfügbaren Business Intelligence-Technologien, steht im Fokus des nachfolgenden Kapitels. Abschließend werden einzelne Themen, denen im Rahmen der IT-Umsetzung eine besondere Bedeutung zukam und deren Problemstellungen typisch für vergleichbare Projekte sein dürften, erläutert.

1 Einführung

In der Unternehmenspraxis ist in den letzten Jahren ein Trend hin zu einem Typ von Berichts-anwendung festzustellen, bei dem Kennzahlen oder auch „Key Performance Indicators“ (KPIs) zusammengefasst und in integrierter Form visuell aufbereitet und dargestellt werden. Für diese Art von Anwendung wird häufig der Begriff „Management Cockpit“ verwendet, der auch im Mittelpunkt dieses Beitrags steht. Stellvertretend finden sich häufig die Begriffe Führungscockpit oder Kennzahlen-Cockpit sowie eine Reihe anderer Termini, wie IT-Dashboard, Cockpit-Chart usw., die zwar weit gefasst sind, jedoch im Grunde alle in dieselbe Richtung zielen (vgl. [Wi08]).

Im engeren Sinne leitet sich der metaphorische Begriff „Management Cockpit“ vom Bild des Cockpits in einem Flugzeug ab, das die verschiedenen Steuerungs- und Anzeigeeinstrumente zur Steuerung des Flugzeugs beinhaltet. Aufbauend auf dem Konzept von *Patrick M. Georges*, der das Management Cockpit ursprünglich als Instrument zur Optimierung der Entscheidungsfindung im Team entwickelt hat (vgl. [Ge00]), hat die SAP AG in den 90er-Jahren im Rahmen des Strategic Enterprise Management (SEM) ein solches Management Cockpit entwickelt. „Es handelt sich dabei um einen eigens entwickelten Raum mit Bildschirmen an den Wänden. Die Grundstruktur des Raumes erinnert an eine Einsatzleitwarte (...), während die Details die Sicht eines Piloten im Cockpit widerspiegeln. Der Grundaufbau ist so gewählt, dass mehrere Manager die Möglichkeit haben, sich gleichzeitig über die kritischen Erfolgsfaktoren zu informieren. Dazu werden diese auf verschiedenen Bildschirmen an den Wänden projiziert. SAP bezeichnet diesen Raum als ‚war room‘, der an ein Flugzeugcockpit erinnern soll (...).“ [Re00]

Im Rahmen des Beitrags wird der Begriff „Management Cockpit“ im in der Praxis geläufigen, weiter gefassten Sinne verwendet. Auch wenn wesentliche Elemente des „echten“ Management Cockpits wie z.B. das räumliche Konzept fehlen, so finden sich auch hier analog zum Flugzeug-Cockpit oft Ampel-, Tachometer- oder Thermometer-Darstellungen, die dazu beitragen sollen, große Mengen von Informationen übersichtlicher darzustellen.

2 Projektbeschreibung

2.1 Projekthistorie

Die Ursprünge des hier beschriebenen Projekts reichen mehrere Jahre zurück. Die ersten Versuche im Unternehmen, die Anforderungen an ein Management Cockpit durch geeignete IT-Werkzeuge zu unterstützen, waren zunächst wenig erfolgreich. Beispielsweise scheiterte die Abbildung der Profit-Center-Rechnung (vgl. 3.1) im Unternehmens-Reporting, das auf dem SAP Business Warehouse (SAP BW) basiert, rückblickend betrachtet u.a. daran, dass zunächst versucht wurde, die komplexe Struktur eines in SAP R/3 bestehenden Profit-Center-Berichts in SAP BW entsprechend nachzubilden. Dies gelang zwar, jedoch erwiesen sich später notwendige Anpassungen, die jeweils parallel

im ERP- sowie BI-System durchgeführt werden mussten, als zu aufwändig und nicht flexibel genug.

Vor dem Hintergrund dieser Erfahrungen wurde seitens des Controlling im Jahr 2007 damit begonnen, ein Management Cockpit, das zentrale Kennzahlen zur Unternehmenssteuerung enthält zu definieren und dieses zunächst in Excel-basierter Form umzusetzen. Im Mai 2007 wurde in einem ersten Treffen zwischen Controlling- und IT-Vertretern eine Projektstudie durchgeführt, mit dem Ziel die Machbarkeit eines Management Cockpits basierend auf SAP BW abzuschätzen. Aufgrund des positiven Ergebnisses hinsichtlich der Machbarkeit wurde mit der Umsetzung des Projekts im Juli 2007 begonnen. Das Projekt wurde unter Beteiligung folgender Projektrollen durchgeführt: eine fachliche Projektleitung aus dem Bereich Unternehmenscontrolling, eine Assistenz der Projektleitung, ein Modulverantwortlicher SAP BW, der gleichzeitig die Projektleitung seitens der IT inne hatte, sowie ein Modulverantwortlicher SAP CO. Bei Bedarf wurden weitere Experten des Fachbereichs zur Unterstützung herangezogen. Die Anwendung wurde vollständig mit SAP Netweaver BI 7.0 realisiert. Die Produktivsetzung des Management Cockpits erfolgte im Februar 2008.

2.2 Projektziele

Im Zuge der Anforderungsdefinition wurden die folgenden allgemeinen Anforderungen an das zu entwickelnde Management Cockpit definiert [Ma08]:

- Früherkennung und vorausschauende Steuerung durch eine differenzierte Darstellung des Plan-/Forecast-Vergleichs aller steuerbaren Einflussgrößen des operativen Ergebnisses (EBIT),
- integrierte Bereitstellung von Kommentaren und Ursachenanalysen,
- hohe Aktualität der Berichterstattung durch eine tägliche Aktualisierung des Datenbestands,
- Reduzierung des Aufwands durch eine weitgehende Automatisierung der Berichterstattung,
- Transparenzsteigerung durch unternehmensweit einheitliche Kennzahlen sowie eindeutig definierte Datenquellen und Berechnungsgrundlagen,
- Integration der Kennzahlen aus verschiedenen Systemen in einer zentralen Anwendung,
- Benutzerfreundlichkeit durch die Darstellung aller relevanten Informationen auf einer Seite sowie die Visualisierung der Unternehmenskennzahlen durch Tachos, Ampeln, etc.

3 Betriebswirtschaftliche Anforderungen

3.1 Steuerungskonzept

Das im Beitrag beschriebene Management Cockpit beruht auf einem im Rahmen des Projekts „NCO – Neuausrichtung Controlling“ entwickelten betriebswirtschaftlichen Steuerungskonzept, das mit der *Segmentsicht* sowie der *Sicht nach Wertschöpfungsstufen* zwei zentrale Ebenen zur Unternehmenssteuerung umfasst [MH04]:

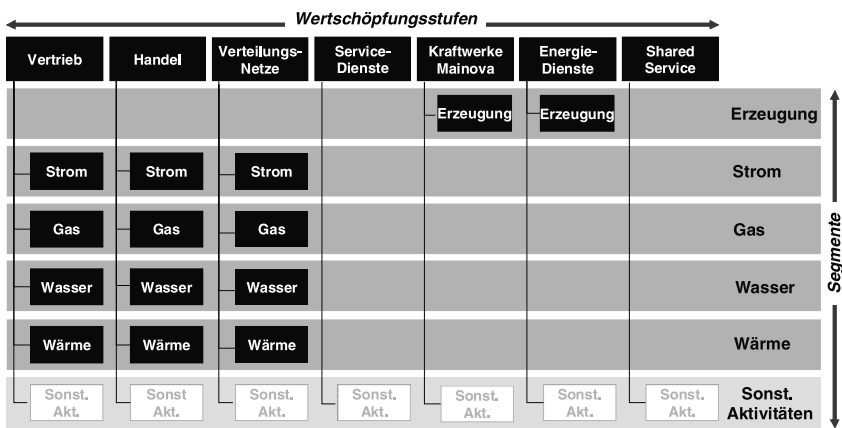


Abbildung 1: Steuerungsebenen Management Cockpit

Während sich die Segmentsicht aus den Geschäftsfeldern der Mainova AG – Erzeugung, Strom, Gas, Wärme, Wasser sowie sonstigen Aktivitäten ergibt, steht hinter der Ebene der Wertschöpfungsstufen ein seit 2004 eingeführtes Profit-Center-Konzept. Innerhalb dieses Konzepts nehmen die einzelnen Wertschöpfungsstufen die Rolle als weitgehend eigenständig agierende und gewinnorientierte Profit-Center ein. Die Profit-Center-Struktur orientiert sich hierbei an der Wertschöpfungskette; von der Erzeugung bzw. dem Energiebezug über die Verteilung bis hin zu Vertrieb und Abrechnung [Ma07a]. Verbunden mit dem Profit-Center-Konzept ist eine Systematik zur internen Leistungsverrechnung, welche die zwischen den Profit-Centern bestehenden Leistungsbeziehungen abbildet. Hierzu wurden für die teilnehmenden Unternehmensbereiche und Gesellschaften im Rahmen von *Service Level Agreements* (SLAs) jeweils eigene „Produkte“ definiert, die es jedem Profit-Center, d.h. auch unternehmensinternen Serviceeinheiten ermöglichen, eigene Erlöse zu erwirtschaften. In engem Zusammenhang mit der Profit-Center-Rechnung steht ein seit 2008 wirksames Verfahren zur Beurteilung und leistungsbezogenen Vergütung der Mitarbeiter, durch das die Ergebnisbeiträge der einzelnen Mitarbeiter mit dem unternehmerischen Erfolg der einzelnen Profit-Center verknüpft werden [Ma07b].

3.2 Definierte Kennzahlen

Im Rahmen des Projekts wurden die folgenden Kennzahlen ausgewählt, mit dem Ziel der Unternehmensführung einen Überblick über die wichtigsten Steuerungsinformationen zu geben [Ma08]:

Zentrale Steuerungsgrößen	
Ergebnis IFRS Mainova Konzern	EBIT IFRS je Segment
Ergebnis HGB Mainova AG	Personalkosten HGB

Betriebs- und Finanzergebnis	Steuerungsgrößen der Unternehmenseinheiten
Rohmarge I Mainova Verbund (IFRS)	Rohmarge I je Wertschöpfungsstufe
Quartalsergebnis Mainova Verbund (IFRS)	
KPL gesamt	KPL je Wertschöpfungsstufe
IPL gesamt	IPL je Wertschöpfungsstufe

Abbildung 2: Zentrale Steuerungsgrößen des Management Cockpits

Zu den wichtigsten Ergebnisgrößen gehören ROCE (Return On Capital Employed) zur Darstellung der Rentabilität des eingesetzten Kapitals, EBT (Earnings Before Taxes, Jahresüberschuss vor Steuern) sowie die durchschnittliche Bilanzsumme. Diese Kennzahlen werden sowohl nach dem internationalen Rechnungslegungsstandard IFRS als auch nach HGB-Vorschriften ausgewiesen. Jede Kennzahl wird immer als Ist-Wert je Quartal und Jahr zusammen mit den entsprechenden Vorjahreswerten sowie als jahresbezogener Plan- und Forecast-Wert dargestellt. Die Kennzahl EBIT (Earnings Before Interest and Taxes, Betriebsergebnis) wird zusätzlich je Segment ausgewiesen. Als weitere Ergebnisgrößen werden die Gewinn- und Steuerumlage sowie die Personalkosten dargestellt.

Zentral für die vorausschauende Steuerung ist die Kennzahl Rohmarge I, mit der die Differenz zwischen Erlösen und direkten variablen Kosten ausgewiesen wird. Die Rohmarge I wird zum einen je Segment und darüber hinaus differenziert nach Wertschöpfungsstufe tagesgenau im Management Cockpit ausgewiesen. Für die Rohmargen wird regelmäßig ein Forecastergebnis gebildet, welches mit dem geplanten Ergebnis verglichen wird. Liegt die Differenz außerhalb eines vorgegebenen Toleranzintervalls, wird dies signalisiert. Das Management wird damit frühzeitig auf außerplanmäßige Abweichungen hingewiesen und kann gegebenenfalls gegensteuern.

Wesentlich für die Unternehmenssteuerung sind weiterhin die Kennzahlen der Investitions- und Kostenprojektplanung. Investitionen und Kostenprojekte werden ab einer bestimmten Größenordnung bei der Mainova AG im Rahmen der jährlich durchgeführten *Maßnahmenplanung* erfasst. Der damit verbundene Prozess wurde bereits vor mehreren

Jahren reorganisiert und unterstützt durch SAP BW abgebildet (vgl. [Ra06]). Im Management Cockpit werden auf Ebene der einzelnen Wertschöpfungsstufen (Vertrieb, Handel, Verteilung, ...) Ist-, Plan- und Forecastwerte aus der Investitions- und Kostenprojektplanung (IPL, KPL) sowie der jeweilige Ausschöpfungsgrad (Ist-Forecast-Vergleich) abgebildet. Weicht der Ausschöpfungsgrad der Kosten- und Investitionsprojekte vom jeweiligen Soll-Ausschöpfungsgrad ab, wird dies per Farbsignal angezeigt.

Sowohl für den Forecast-/Planvergleich der Rohmargen wie auch für den Vergleich der Ausschöpfung der Kosten- und Investitionsprojekte im Soll und im Ist gilt, dass die Toleranzbereiche für die Einzelergebnisse individuell einstellbar sind. So kann bei Bedarf die Signalgebung bei speziellen Geschäftsbereichen besonders sensibel justiert werden, so dass bei Abweichung frühzeitig gewarnt wird.

Aus dem Betriebs- und Finanzergebnis nach IFRS auf Ebene des Unternehmensverbunds stammen weiterhin die Ergebnisse Abschreibungen, Beteiligungen, Finanzerträge, Finanzaufwendungen sowie EBT für die, wie auch für die Rohmarge I, ein Vorjahresvergleich auf Ebene der einzelnen Segmente durchgeführt wird.

4 Informationstechnische Umsetzung

4.1 Vorgehen im Projekt

Im Rahmen des Projektes wurde kein explizites Vorgehensmodell im Sinne einer durchgängigen Methode angewendet. Es lassen sich aber zumindest einzelne Projektschritte vor dem Hintergrund eines für BI-Projekte typischen Vorgehens (vgl. [Ke06], S.147ff) einordnen. Die folgenden Schritte waren für das Projektvorgehen kennzeichnend:

- Definition und Aufnahme der fachlichen Anforderungen (auf einer nicht zu detaillierten Ebene),
- Bestimmung von Kern-Themen, die im Rahmen des Projektes zu lösen sind (z.B. Plandateneingabe, Abbildung Profit-Center-Rechnung),
- Identifikation der Datenquellen (Welche Daten sind bereits in SAP BW vorhanden? Welche Daten müssen neu beschafft werden? Was sind die relevanten Datenquellen? Für welche Daten muss eine Plandateneingabe realisiert werden?),
- Konzeption des Datenmodells sowie der Architektur (vgl. 4.2),
- Aufbau der erforderlichen Strukturen in SAP R/3 (Profit-Center-Gruppen, Kostenarten-Gruppe),
- Definition und Modellierung der Kennzahlen in SAP R/3 und SAP BW sowie anschließender Abgleich der Kennzahlen zwischen SAP R/3 und SAP BW,
- Umsetzung des Datenmodells und der IT-Architektur (Aufbau InfoCube für Plandatenerfassung, Aufbau MultiProvider, Erstellung der Queries und der Web Templates, ...),

- Test und Überführung in den Produktivbetrieb.

Während der gesamten Projektdauer fanden regelmäßige Treffen des Projektteams statt, in denen die Projektbeteiligten ihr Feedback zum jeweiligen Entwicklungsstand einbringen konnten. Die bereits zu einem frühen Zeitpunkt zur Verfügung stehende Oberfläche des Management Cockpits wurde anhand des Feedbacks sukzessive weiterentwickelt; im Verlauf dieses für BI-Projekte typischen Vorgehens (vgl. [Ke06], S. 164ff) wurde die Anwendung ausgehend vom Prototypen in die abschließende produktive Anwendung überführt.

4.2 IT-Architektur

Die dem Management Cockpit zugrunde liegende technische Architektur lässt sich wie folgt in verschiedenen Schichten darstellen (vgl. Abb. 3):

- Transaktionsebene (SAP R/3): Im eingesetzten OLTP-System SAP R/3 finden die eigentlichen Buchungen auf den hierfür vorgesehenen Kontierungselementen (Aufträge, PSP-Elemente u.a.) statt. Für das Management Cockpit relevant sind zum einen die im Rahmen der unternehmensinternen Maßnahmenplanung durchgeführten Kontierungen zu den im SAP-Modul PS erfassten Investitions- und Kostenprojekten zusammen mit ihren Stammdaten (vgl. [Ra06]).

Zum anderen handelt es sich dabei um die im Rahmen der Profit-Center-Rechnung erfassten statistischen Buchungen, die auf einem relevanten Teil der aus verschiedenen R/3-Modulen stammenden Bewegungsdaten basieren (vgl. [Sa02]). Darüber hinaus sind die im Rahmen des Projekts definierten und im R/3-System hinterlegten Profit-Center-Gruppen hervorzuheben, mit Hilfe derer die beiden zentralen Ebenen zur Unternehmenssteuerung abgebildet sind (vgl. 3.1). Die auf der Transaktionsebene stattfindenden Änderungen werden mit Hilfe einer für diesen Zweck eingerichteten Prozesskette tagesaktuell in das SAP BW übernommen.

- Datenhaltungsebene: Die innerhalb des SAP BW angesiedelte Datenhaltungsebene des Management Cockpits besteht aus insgesamt vier verschiedenen, mehrdimensionalen InfoCubes mit den zugehörigen Bewegungs- und Stammdaten sowie der vorgelagerten Persistent Staging Area (PSA). Zu den Datenzielen zählen je ein InfoCube für die Plan- sowie die Ist-Daten der Maßnahmenplanung, ein InfoCube für die Daten der Profit-Center-Rechnung sowie ein weiterer beplanbarer (transaktionaler bzw. realtime-fähiger) InfoCube (vgl. [Eg06], S. 458), der zur Aufnahme quartals- und jahresbezogener Planwerte, Forecastwerte und Abweichungsparameter sowie verschiedener Kennzahlen dient, die derzeit nicht unmittelbar aus dem SAP-Quellsystem gewonnen werden können.¹

¹ Aus unterschiedlichen Gründen ist für viele Kennzahlen (noch) eine manuelle Erfassung notwendig (vgl. [Oe06], S. 53).

Als Grundlage für die Plandateneingabe dienen zwei sogenannte Aggregations-ebenen [Sa07], welche die Granularität der Planung auf Jahres- und Quartals-ebene festlegen.

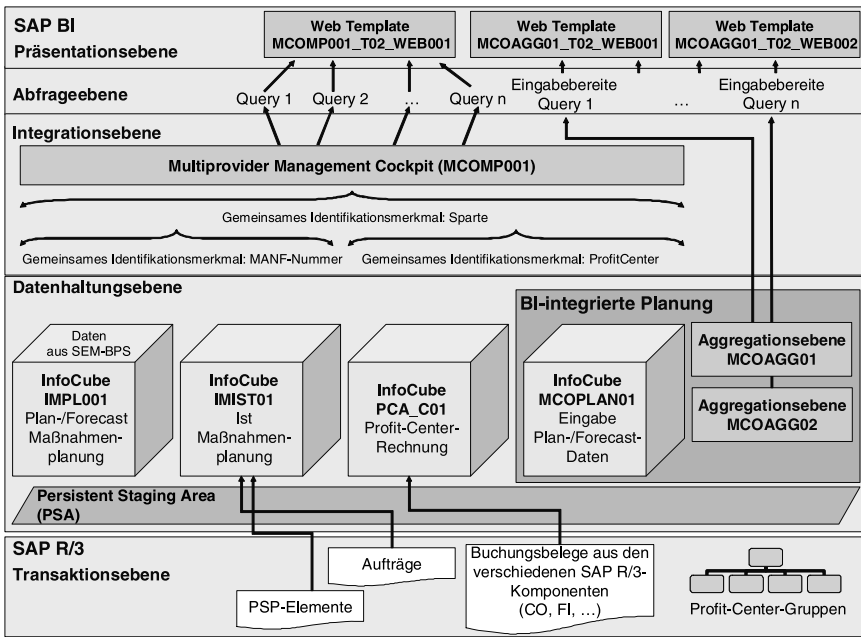


Abbildung 3: IT-Architektur

- **Integrationsebene:** In dieser Ebene werden die einzelnen InfoCubes der Datenhaltungsebene zu einem sogenannten MultiProvider zusammengeführt und gemeinsam für die Datenanalyse zur Verfügung gestellt. Der MultiProvider enthält selbst keine Daten; seine Daten ergeben sich ausschließlich aus den zugrunde liegenden InfoProvidern, die per Union-Operation zusammengefasst werden. Bei der Modellierung des Multiproviders spielt die Identifikation von Merkmalen eine zentrale Rolle, da jedes Merkmal des MultiProviders eindeutig einem Merkmal eines oder mehreren der beteiligten InfoProvider zugeordnet werden muss [Sa07].
- **Abfrageebene:** Innerhalb dieser Ebene sind die einzelnen Queries definiert, mit denen die Daten auf Basis des MultiProviders abgefragt werden. Um die Kennzahlen aus den verschiedenen Datenquellen innerhalb der Abfragen sachlich korrekt gegenüberzustellen, werden die gemeinsamen Merkmale und Navigationsattribute bzw. sich entsprechende Merkmale genutzt, um die Kennzahlen auf der Grundlage der durch die Wertschöpfungs- und Segmentsicht gekennzeichneten Struktur des Management Cockpits abzubilden. Innerhalb der Queries sind außerdem die verschiedenen Schwellenwerte (Exceptions) definiert, um die für die Anwendung geforderte Ampelfunktionalität zur Unterstützung der Erkennung von Abweichungen abzubilden. Einen weiteren Bestandteil der Abfrageebene bilden die sogenannten „eingabebereiten Queries“, die zur Eingabe

der für das Management Cockpit benötigten Plan- und Forecastdaten sowie manuell erfasster Kennzahlen dienen.

- Präsentationsebene: In dieser Ebene werden die einzelnen Queries unter einer einheitlichen Benutzeroberfläche zusammengefasst. Für die einzelnen Abfrageelemente (Kennzahlen, Exceptions, ...) werden Komponenten (Tabellen, Charts, ...) sowie Einstellungen (z.B. Farben und Symbole für Schwellenwert-Unter-/Überschreitungen, ...) definiert, mit denen sich das Erscheinungsbild der Anwendung gestalten lässt. Technisch sind alle für das Management Cockpit relevanten Abfragen in einem „Web Template“ zusammengefasst. Für die Eingabe der jahres- und quartalsbezogenen Plan- und Forecastwerte existiert je ein eigenes Web Template, das die eingabebereiten Queries enthält. Die Übergabe von erforderlichen Variablenwerten geschieht für den Anwender transparent: Eine Navigationsoberfläche als Bestandteil der Präsentationsebene ermöglicht die Auswahl zentraler Selektionsparameter (wie z.B. den Abfragezeitraum) sowie den möglichst einfach gestalteten Aufruf des Management Cockpits.

4.3 Ausgewählte Implementierungsaspekte

Ein zentrales Thema im Rahmen der Implementierung war die Abbildung der „Kennzahlen-Matrix“, die sich aus den beiden maßgeblichen Strukturierungsebenen nach *Segmenten* sowie nach *Wertschöpfungsstufen* ergibt (vgl. 3.1). Um die beiden Ebenen abzubilden, wurden in SAP R/3 zwei Profit-Center-Gruppen (Hierarchien) angelegt, die die Profit-Center (bzw. untergeordnete Profit-Center-Gruppen) zum einen dem jeweiligen Segment und zum anderen der jeweiligen Wertschöpfungsstufe zuordnen:

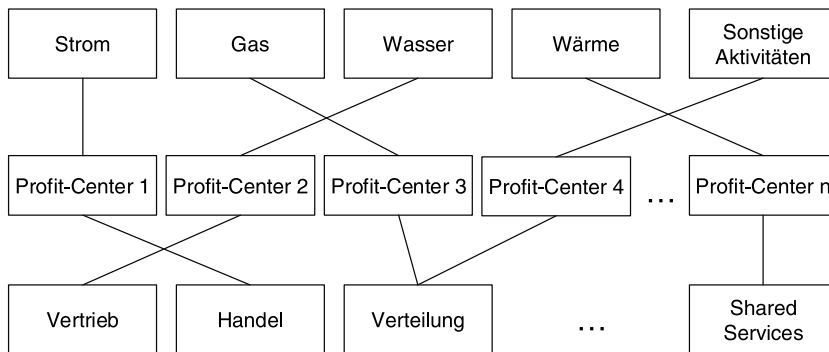


Abbildung 4: Alternative Hierarchien zum Profit-Center

Mit Hilfe dieser beiden Hierarchien können die aus der Profit-Center-Rechnung stammenden Kennzahlen sowohl gemäß der Segmentsicht als auch gemäß der Zugehörigkeit zur Wertschöpfungsstufe dargestellt werden. Während die nach Segmenten strukturierte Profit-Center-Gruppe primär dazu dient, die Gesamtwerte je Segment für einzelne Kennzahlen der Profit-Center-Rechnung zu ermitteln, wird die alternative nach Wertschöpfungsstufen strukturierte Hierarchie genutzt, die Kennzahlen je Wertschöpfungsstufe abzubilden.

Für Wertschöpfungsstufen, bei denen eine segmentbezogene Betrachtung gefordert ist, ist diese Hierarchie nochmals weiter nach Segmenten untergliedert, damit auch Kennzahlen wie z.B. „Rohmarge Vertrieb Strom“ oder „Rohmarge Vertrieb Gas“ dargestellt werden können.²

In diesem Zusammenhang ist bemerkenswert, dass der zuerst angestrebte Versuch, Kennzahlen für die „Schnittpunkte“, die sich aus der Kombination Segment und Wertschöpfungsstufe ergeben, direkt mit Hilfe der beiden definierten alternativen Hierarchien abzubilden, nicht das gewünschte Ergebnis erbrachte. Die Grundidee hinter diesem ursprünglichen Lösungsansatz war, die Segmentinformation aus der nach Segmenten strukturierten Profit-Center-Hierarchie zu lesen und diese in das Attribut „Sparte“ des Merkmals Profit-Center fortzuschreiben. Eine Kennzahl wie z.B. „Rohmarge Vertrieb Strom“ wäre demnach durch eine Merkmalseinschränkung auf die Ausprägung „Strom“ des Profit-Center-Attributs Sparte zusammen mit einer Einschränkung des Profit-Centers auf die Wertschöpfungsstufe „Vertrieb“ abzubilden gewesen. Wie sich dann aber im Projekt zeigte, stimmten die Werte für auf diese Weise definierte Kennzahlen in bestimmten Fällen nicht mit denen des Quellsystems SAP R/3 überein. Dies war immer genau dann der Fall, wenn ein Eliminierungs-Profit-Center, das bei der Ausführung der Binnenumsatzeliminierung herangezogen wird, einem anderen Segment zugeordnet war als das mit dem Eliminierungs-Profit-Center in Beziehung stehende Profit-Center selbst. Die Einschränkung auf ein Segment in Kombination mit der Binnenumsatzeliminierung führte dazu, dass die Kennzahlen mit diesem zunächst elegant erscheinenden Lösungsansatz nicht korrekt abgebildet werden konnten. Mit Hilfe der Abbildung der Segmente als Teilhierarchien innerhalb der Wertschöpfungsstufen-Hierarchie ließ sich dieses Problem jedoch lösen.³

Die Gegenüberstellung der *Kennzahlen aus der Investitions- und Kostenprojektplanung* nach denselben Strukturierungsebenen basiert hingegen zum einen auf dem zu jeder geplanten Maßnahme gepflegten Stammdaten-Attribut „Sparte“ sowie zum anderen auf der ebenfalls nach Wertschöpfungsstufen organisierten Programmpositions-Hierarchie. Aufgrund der Verwendung eines Multiproviders können die in unterschiedlichen Info-Cubes vorgehaltenen Kennzahlen der Profit-Center-Rechnung sowie der Investitions- und Kostenprojektplanung gemeinsam innerhalb einer Query dargestellt werden.

Die *Modellierung der Kennzahlen* wurde zunächst ohne Bezug zu den oben genannten Hierarchien vorgenommen. Eine Einschränkung der Kennzahlen auf die jeweilige Wertschöpfungsstufe bzw. das jeweilige Segment erfolgte in einem zweiten Schritt im Rahmen der Query-Definition. Als wesentliches, die jeweilige Kennzahl bestimmendes Merkmal ist für die Profit-Center-Rechnung vor allem die *Kostenart* zu nennen, für die ebenfalls eine eigene Kostenarten-Gruppe definiert wurde. Um die Kennzahlen in Hin-

² Vorteilhaft in diesem Zusammenhang ist, dass Kennzahlen, die den Schnittpunkt zweier Strukturen darstellen, im SAP BW Query Designer zellenweise definiert werden können. Der Wert solcher „Ausnahmezellen“ übersteuert die implizit erzeugten Werte der Query (vgl. [Sa07]).

³ Auf Probleme und Lösungsansätze beim Einsatz alternativer Hierarchien geht z.B. K. Oehler ein (vgl. [Oe06], S. 140ff).

blick auf die zu späteren Zeitpunkten erforderlichen Anpassungen zentral an einer Stelle pflegen zu können, wurden sie als global gültige (wiederverwendbare) Kennzahlen angelegt, die in allen Queries zum InfoProvider verwendet werden können (vgl. [Sa07]).

Im Rahmen der Profit-Center-Rechnung war zudem gefordert, die in SAP R/3 bereits im Standard vorhandene *Binnenumsatzeliminierung* nachzubilden. Hierzu wurden in SAP BW sogenannte „Kennzahlen mit Referenz“ angelegt (vgl. [Sa07]). Als Merkmalspaar für die Binnenumsatzeliminierung wurden für diese Kennzahlen jeweils das Merkmal Profit-Center sowie das Merkmal Eliminierungs-Profit-Center angegeben. Die Verwendung der Referenz-Kennzahlen führt dazu, dass Binnenumsätze zwischen zwei Profit-Centern auf Ebene übergeordneter Hierarchieknoten unterdrückt werden.

Aus unterschiedlichen Gründen können derzeit nicht alle Kennzahlen direkt aus ihren Quellsystemen in das Management Cockpit übernommen werden. Zudem werden bestimmte Plan-, Forecast- und Abweichungswerte ausschließlich innerhalb des Management Cockpits erfasst. Zur Erfassung dieser Daten wurde ein eigener InfoCube definiert (vgl. 4.2). Die Datenerfassung findet dabei über die im Rahmen der BI-integrierten Planung mit SAP BI 7.0 neu hinzugekommenen Funktionalität der „eingabebereiten Queries“ statt. Hierbei wird insbesondere von der Möglichkeit Gebrauch gemacht, Daten auch auf Ebene von Hierarchieknoten, d.h. als quasi aggregierte Werte, eingeben zu können. Technisch wird dabei ein erfasster Wert gemäß einer vorgegebenen Verteilungsart durch Disaggregation auf die darunter liegenden Detailebenen verteilt (vgl. [Sa07]). Durch die Anwendung dieser Technik können die jahres- bzw. quartalsbezogenen Plan- und Forecastwerte auf Ebene der für das Management Cockpit maßgeblichen Steuerungsebenen (Segmentsicht und Sicht nach Wertschöpfungsstufen) erfasst und den aus der Profit-Center-Rechnung stammenden Istwerten gegenüber gestellt werden.

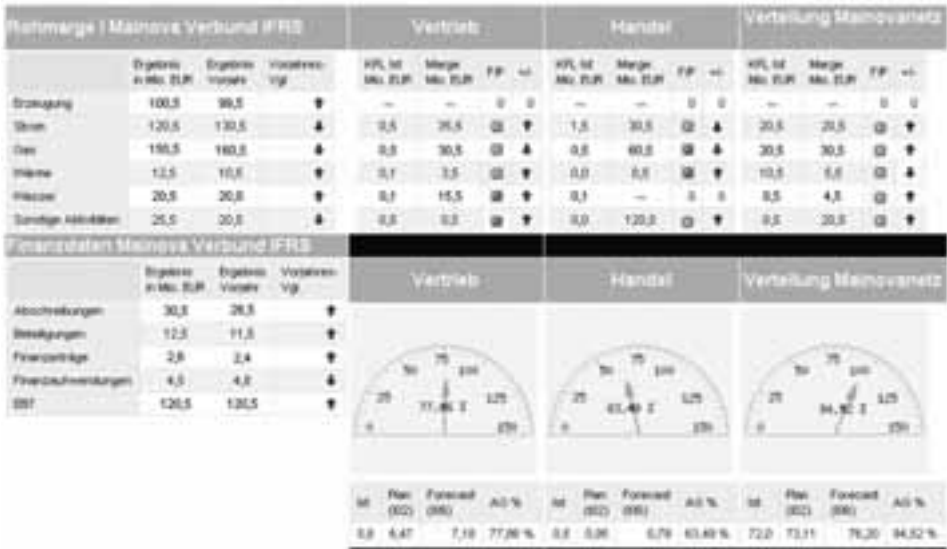


Abbildung 5: Management Cockpit (Ausschnitt)

Eine wesentliche Anforderung bestand nicht zuletzt darin, die Visualisierung der Unternehmenskennzahlen durch Tachos, Ampeln, u.a. zu unterstützen (vgl. 2.2).

Die folgenden Abweichungsgrößen wurden als Basis für das darauf aufbauende Exception Reporting definiert:

- Absolute Abweichung aktueller Wert minus Vorjahreswert (für Kennzahlen der Profit-Center-Rechnung),
- absolute Abweichung Forecastwert im Verhältnis zum Planwert (für die Kennzahl Rohmarge I),
- Forecastwert innerhalb eines festgelegten Intervalls relativ zum Planwert (für die Kennzahl Rohmarge I),
- Ausschöpfungsgrad des Investitions- und Kostenprojektplans in Prozent (Ist/Forecastvergleich).

Die Schwellenwerte bei denen eine Abweichung im Management Cockpit positiv oder negativ dargestellt wird (z.B. durch grüne/rote Symbole), sind derzeit fest hinterlegt und werden quartalsweise angepasst. Dies ist erforderlich, da ein Ausschöpfungsgrad von 95% für den Investitions- und Kostenprojektplan im vierten Quartal eine andere Bedeutung hat als im ersten Quartal.⁴ Das die relative Forecast-/Planabweichung der Rohmarge I bestimmende Intervall kann über die Eingabe eines Parameters für jedes Berichtsquartal angepasst werden.

5 Zusammenfassung und Ausblick

Im Vordergrund des in diesem Beitrag beschriebenen Projekts stand das Bestreben, Kennzahlen, die auf Daten aus unterschiedlichen Quellen beruhen, unter einer einheitlichen Benutzeroberfläche in integrierter und visuell aufbereiteter Form im Sinne eines Management Cockpits darzustellen. Eine zentrale Herausforderung bestand darin, die aus unterschiedlichen Quellen stammenden Kennzahlen entlang der beiden maßgeblichen Strukturierungsebenen nach Segmenten sowie nach Wertschöpfungsstufen abzubilden (vgl. 3.1).

Eine zentrale Frage, die auch schon bei früheren Konzepten der Führungsinformationssysteme (Executive Information Systems u.a.) vernachlässigt wurde, ist die Frage, woher die Daten bezogen werden sollen (vgl. [Oe06], S. 17ff). Im Rahmen des Projektvorgehens (vgl. 4.1) war deshalb zu klären, welche Daten bereits im Datawarehouse vorhanden sind und welche erst noch aus Quellsystemen beschafft werden müssen. Der Projektaufwand steigt mit den für die Datenbeschaffung erforderlichen Aktivitäten i.d.R. erheblich an (vgl. [Ke06], S. 169). Im Projekt wurden u.a. die Daten der Profit-Center-Rechnung erstmalig in das SAP BW übernommen. Durch die Verwendung

⁴ Künftig sollen die Schwellenwerte dynamisch, basierend auf dem erwarteten durchschnittlichen Kostenverlauf, an den jeweiligen Berichtszeitpunkt angepasst werden.

des sogenannten *Business Content* (vorgefertigte ETL-Prozesse, Datenmodelle, usw.), der mit SAP BW ausgeliefert wird, konnte der resultierende Aufwand für die Implementierung jedoch stark reduziert werden.

Für die Daten, die bereits im BI-System vorhanden waren, stellte sich im Rahmen des Projekts zudem die Frage, wie die Daten gemäß den fachlichen Vorgaben in integrierter Form dargestellt werden können. Auch dieser Punkt wird in entsprechenden Projekten häufig vernachlässigt (vgl. [Oe06], S. 44). Um beispielsweise ein Ausweisen von Kennzahlen auf Ebene von Wertschöpfungsstufen zu erreichen, muss die Information „Zugehörigkeit zur Wertschöpfungsstufe“ in irgendeiner Form vorhanden sein. Im Projekt wurden diesbezüglich im Quellsystem SAP R/3 entsprechende Profit-Center-Gruppen eingerichtet, welche die einzelnen Profit-Center den jeweiligen Wertschöpfungsstufen zuordnen.

Erst nachdem diese strukturellen Voraussetzungen geschaffen sind, ist es möglich, eine Datenintegration unterschiedlicher Quellen im engeren Sinne zu erreichen, die über ein – nicht wirklich integriertes – „Nebeneinanderstellen“ von Kennzahlen unter einer gemeinsamen Oberfläche hinaus geht. Im Projekt wurde für die anschließende Zusammenführung der Datenquellen zu einer gemeinsamen Datenanalyse-Schicht das *MultiProvider-Konzept* verwendet, durch das auch die *Erweiterbarkeit* der Anwendung unterstützt wird (vgl. [Sa07]). Die gemeinsamen Merkmale (z.B. das Merkmal Sparte) müssen dabei zum einen für alle relevanten InfoProvider als Identifikationsmerkmale bei der MultiProvider-Definition mit angegeben werden, zum anderen muss im Rahmen der Query-Definition eine entsprechende Selektion anhand dieser Merkmale erfolgen.

Die im Beitrag beschriebene Vorgeschichte des Projekts (vgl. 2.1) macht deutlich, dass ein Ansatz, der die erforderlichen Strukturen für das auf SAP BW basierende Reporting lediglich auf Basis komplexer definierter Regeln nachbildet, zu Problemen bei späteren Anpassungen führen kann. Vielmehr muss bereits während der Entwurfsphase daran gedacht werden, die Strukturen so flexibel zu gestalten, dass spätere Anpassungen einfach möglich sind und im Idealfall automatisch zwischen dem Quellsystem und dem Datawarehouse abgeglichen werden, um ein späteres „Auseinanderlaufen“ von Auswertungsergebnissen zu vermeiden. Innerhalb des Projekts waren die definierten *Hierarchien* hierfür das adäquate Mittel.

Im Rahmen der Entwicklung der Anwendung wurde zudem darauf geachtet, die Modellierung der Kennzahlen von den durch die Organisation vorgegebenen Strukturen (Segmente, Wertschöpfungsstufen) zu trennen. Die definierten Kennzahlen sind zunächst ohne Bezug zu den Organisationsstrukturen als globale Kennzahlen auf InfoProvider-Ebene hinterlegt. Erst bei der Übernahme einer Kennzahl in einzelne Queries wird die Kennzahl in Beziehung zu den Wertschöpfungsstufen bzw. Segmenten gesetzt. Auf diese Weise wird die *Wiederverwendbarkeit* und *Anpassbarkeit* der Kennzahlen unterstützt. Durch die Verwendung einer geeigneten Hierarchie für die Kostenarten sind Änderungen der Kennzahl-Definitionen ebenfalls einfach und an zentraler Stelle zu bewerkstelligen.

Das Management Cockpit in seiner jetzigen Form beschränkt sich auf die Darstellung der Kennzahlen für die definierten Aggregationsebenen. Ein *Drill-Down* auf die dahinter liegenden Detaildaten ist bisher nicht möglich. Alternativ dazu besteht jedoch die Möglichkeit, über die in das Management Cockpit integrierte *Dokumentationsfunktion* die dort hinterlegten detaillierten Segmentberichte sowie Kommentare usw. abzurufen. Für den weiteren Ausbau des Management Cockpits ist u.a. eine Drill-Down-Möglichkeit für verschiedene Kennzahlen sowie eine weiter automatisierte Bereitstellung der derzeit teils noch manuell erfassten Daten geplant.

6 Literaturverzeichnis

- [Eg06] Egger, N. et al.: SAP Business Intelligence, 1. Aufl., Bonn 2006.
- [Ge00] Georges, P. M.: Das Management Cockpit – Benutzungsschnittstelle für Management Software: Ein Erfahrungsbericht, In: Wirtschaftsinformatik 42 (2000) 2.
- [Ke06] Kemper, H.-G. et al.: Business Intelligence – Grundlagen und praktische Anwendungen: Eine Einführung in die IT basierte Managementunterstützung, 2. Aufl., Wiesbaden 2006.
- [Ma07a] Mainova AG: Unternehmenshandbuch, Handbuch Controlling – Profit Center Rechnung, 2007.
- [Ma07b] Mainova AG: Betriebsvereinbarung Nr. 03/2007, Mainova Vario Plus – Leistungsorientierte Vergütung für tarifliche Mitarbeiterinnen und Mitarbeiter auf Grundlage einer Leistungsbeurteilung, 2007.
- [Ma08] Mainova AG: Interne Präsentationsunterlagen zum Management Cockpit, 2008.
- [MH04] Mainova AG/Horváth & Partners Management Consultants: Interne Unterlagen zum Projekt „Neuausrichtung Controlling (NCO), 2004.
- [Oe06] Oehler, K.: Corporate Performance Management mit Business Intelligence Werkzeugen, München Wien 2006.
- [Ra06] Raimann, J. et al.: Integrierte Investitions- und Kostenprojektplanung mit SAP SEM-BPS am Beispiel der Mainova AG. In: Schelp, J. et al. (Hrsg.): Integration, Informationslogistik und Architektur, DW2006, 21.–22. Sept. 2006, Friedrichshafen, 13-34 (2006), Lect. Notes Inform., Proc. (Vol. 90), S. 107-124.
- [Re00] Reiterer, H. et al.: Visualisierung von entscheidungsrelevanten Daten für das Management. In: HMD, Praxis der Wirtschaftsinformatik, Heft 212 04/2000, S. 71-83.
- [Sa02] SAP AG: Profit-Center-Rechnung (EC-PCA). In: SAP-Bibliothek, SAP Standard Release 4.6C, April 2001, 2002.
- [Sa07] SAP AG: SAP Bibliothek, SAP NetWeaver® 7.0 Support Package Stack 14 including BI Content Add-On 3 SP 08, Dezember 2007.
- [Wi08] Wikipedia: <http://de.wikipedia.org/wiki/Kennzahlen-Cockpit>, Stand: Februar 2008.

Das Ende der Magersucht in der Managementinformation

Nicolas Bissantz, Roland Zimmermann

Bissantz & Company GmbH
90409 Nürnberg
{nicolas.bissantz | roland.zimmermann}@bissantz.de

Abstract: Berichte berichten nichts. Das gilt zumindest für viele Managementreports. Die Gründe sind unterschiedlich, häufig hängen die Probleme mit zu geringer Informationsdichte und einem Mangel an Automation zusammen. Wir beschreiben Ansätze und Methoden, die dazu geeignet sind, die Qualität des Berichtswesens deutlich zu steigern.

1 Inhaltlich eindimensionale Berichte in 3D

Über ein Jahrzehnt erfolgreicher Software- und Beratungsprojekte verschafften uns die Gelegenheit, mit Hunderten von Betrieben im Bereich der Datenanalyse zusammenzuarbeiten. Dabei gewannen wir Einblicke in so unterschiedliche Gebiete wie Controlling, Marketing, Marktforschung, Hafenlogistik, Immobilienportfoliosteuerung und Verbrechensbekämpfung [Bis00]. Üblich sind fast immer isolierte Betrachtungen einzelner Fragestellungen. Sie münden in Berichten, die unverbunden nebeneinander stehen, dafür aber oftmals optisch opulent und effektheischend daherkommen. Für die fundierte Analyse sind jedoch mehrstufige Auswertungen notwendig, die auf Frageketten basieren. Erst daraus entstehen Erkenntnispfade, wie sie für die Führung von Betrieben der Privatwirtschaft, des Öffentlichen Dienstes und der Öffentlichen Verwaltung nötig sind. Solche Erkenntnispfade verdienen es, in Berichten dargestellt zu werden.

In der Realität sieht es leider oft anders aus: Als die Landesbank Sachsen Ende 2007 infolge der Hypothekenkrise in Not geriet und in einer Nacht-und-Nebel-Aktion an die Landesbank Baden-Württemberg verkauft wurde, berief sich einer der Verwaltungsräte auf die simplen Ampeldarstellungen der ihm vorliegenden Risikoberichte:

„Man kriegt ja regelmäßig den Risikobericht, da kann man nachgucken. Und da hat man so eine kleine Ampel drin, mit den drei Farben, rot, gelb, grün. Und wenn es grün ist, und der überwiegende Teil war eben im grünen Bereich, hoch geratet von internationalen Agenturen, ‚Triple A‘ heißt das, sagt man, na, dann scheint es so zu gehen.“ Ronald Weckesser, Verwaltungsrat Sachsen LB in den Tagesthemen, 27. August 2007

Ampeldarstellungen, PowerPoint-Präsentationen, 3D-Datengebirgen und der Idee, Managern mit Tachos suggerieren zu wollen, Management sei so spannend wie Porsche

fahren, wohnt dieselbe fatale Idee inne: Die Steuerung von Unternehmen ist ein Kinderspiel, wenn man nur die richtige Software mit einer fetzigen Darstellung nutzt.

„Those gauges, meters and traffic lights are so damn flashy! You can imagine that you're sitting behind the wheel of a German-engineered sports car, feeling the wind whip through your hair as you tear around curves on the autobahn at high speeds, all without leaving your desk.“ [Few06, S. 4]

Gemeinsam ist vielen neueren Ansätzen ihr Mangel an Informationsdichte (gemessen als Verhältnis von Datenwerten pro Flächeneinheit). Wenn Führungskräfte auf dem Weg in die Arbeit den Börsen- oder Sportteil ihrer Tageszeitung studieren, überschreiten sie in der Regel den Zenit an Informationsdichte, der ihnen an diesem Arbeitstag zugemutet wird. Der Börsenteil wartet mit etwa 1.500 Zahlen pro Seite DIN A4 auf, Sporttabellen immer noch mit rund 1.000. Die typischerweise heute propagierten Dashboards arbeiten mit Dichten zwischen 5 und 14 % dieser Werte [Bis06].

Manager entscheiden unter Unsicherheit und mit Risiko. Beides kann durch Information in einem gewissen Ausmaß verringert werden. Je dünner die bereitgestellten Informationen sind, desto schwieriger die Entscheidung. Dass Manager sogar überhaupt nur auf Ausnahmesituationen zu reagieren hätten und sich bei grünen Signalen ansonsten entspannt zurücklehnen können, lehnte Szyperski schon 1978 mit Recht ab:

„Festgeschriebene Schwellenwerte, verbunden mit der Management-by-Exception-Fiktion, sind daher gefährlich. Das Management sollte neugierig sein, d. h. neue Informationsverknüpfungen suchen und nicht nur wie eine Kontrollperson auf einer Schaltbühne aufmerksam dösen“, zitiert nach [MeG02, S. 2]

Fälle wie der Zusammenbruch der Landesbank Sachsen machen deutlich, dass Datenanalysten und Informationsdesigner Verantwortung tragen. Wer grüne Ampeln liefert statt der detaillierten Daten dahinter, leistet keine Entscheidungshilfe, er nimmt die Entscheidung vorweg. Wissentlich oder nicht, der Reportersteller hat für den Manager entschieden. Er ist in seine Rolle geschlüpft.

2 Damit Berichte wieder berichten

Wie gutes Managementreporting auszusehen hat, erkennt man am besten daran, wie es nicht sein sollte. Die Mängel, die wir beobachten können, sind zum Teil handwerklicher Natur, zum Teil beruhen sie auf Fehleinschätzungen. Das Ergebnis ist in beiden Fällen das gleiche. Viele Berichte berichten nichts, sie werden nicht gelesen oder sie werden nicht verstanden. Beides ist vergleichsweise harmlos. Schlimmer sind Fälle, in denen Berichte Vorschub leisten für Fehlentscheidungen des Managements.

2.1 Zeit ist knapp – Zeitverschwendung verhindern

Das Standardargument für die modische Magersucht im Managementberichtswesen ist, dass die Zeit der Führungskräfte knapp bemessen ist. „Ein Managementbericht wird durchschnittlich acht Sekunden studiert.“ Das ist angesichts der Qualität vieler Berichte überraschend viel. Wirft ein Börsianer denn nur einen flüchtigen Blick auf den Börsenteil? Die Zeit, die einer Sache zu widmen ist, sollte eine Funktion ihres Werts sein.

Wenn die Zeit der Empfänger knapp ist, und das ist sie zweifellos, dann sollten Berichte in Einstein'scher Logik so knapp als möglich sein, aber nicht knapper. Das heißt nicht immer, dass sie kurz sein müssen. Sie müssen die Zeit wert sein, die zum Lesen und Verstehen benötigt wird. Weglassen ist keine Option. Wenn das Verknappen nicht verlustfrei gelingt, ist die Darstellungsweise falsch.

Eine hilfreiche Annahme ist, dass unsere Leser intelligenter als wir selbst und sehr beschäftigt sind. Anderenfalls wären wir vielleicht selbst derjenige, der sich den Luxus erlauben kann, sich berichten zu lassen. Davon auszugehen, dass Vorstände nichts von Berichten verstehen, weil sie wenig von elektronischen Berichten halten, wäre jedenfalls voreilig.

Müssen Legenden dechiffriert werden, sind Beschriftungen mehrdeutig, unleserlich oder unvollständig, verpufft wertvolle Empfangskapazität nutzlos. Die verfügbare Zeit wurde schlecht genutzt.

2.2 Raum ist knapp – Schluss mit pathetischen Berichten

Nicht nur die Zeit ist knapp, sondern auch der Raum für unsere Darstellungen. Wir verstehen am besten, was wir innerhalb der Augenspanne überblicken können. Am Bildschirm ist das nicht viel. Die Auflösung auf Papier ist trotz verbesserter Bildschirmtechnologie immer noch deutlich höher. Aber auch auf Papier stehen uns meist nur Blätter der Größe DIN A4, manchmal DIN A3 für unsere Berichte zur Verfügung. Haben wir beide Medien, müssen wir sie entsprechend nutzen. Ein Bericht, der für den Bildschirm optimiert ist, ist auf Papier zu mager.

Noch knapper wird der Raum, wenn unsere Berichte projiziert werden. Die Tendenz, das Berichtswesen an den technischen Anforderungen eines Folienstapels zu orientieren, wird genau aus diesem Grund massiv kritisiert. Der knappe Raum verführt zu Auslassungen, Vereinfachungen und einer hektischen, holzschnittartigen Informationskultur. Zu dieser Auffassung gelangte auch die offizielle Untersuchungskommission, die den Absturz der Raumfähre Columbia aufzuklären hatte. Bei der NASA fanden Risikoevaluationen nicht mehr mittels technischer Berichte, sondern in PowerPoint-Präsentationen statt [Tuf06]. Projektionsbedingungen diktierten Vereinfachung und Verdichtung. Auf dem Weg durch die Organisation gingen entscheidende Informationen verloren. Risiken wurden nicht erkannt oder falsch eingeschätzt.

Nicht immer führt schlechtes Berichten zu Katastrophen, aber: Wenn wir die ohnehin knappe Augenspanne mit unnötigen Linien, Balken, Schattierungen, Rahmen, Skalen und Mustern statt mit Daten füllen, bleibt Darstellungskapazität ungenutzt. Unglücklicherweise kommt derlei aber frei Haus und unaufgefordert. Excel, das weltweit wohl am meisten genutzte Programm für Datendarstellungen, produziert geradezu schwülstige Grafiken.

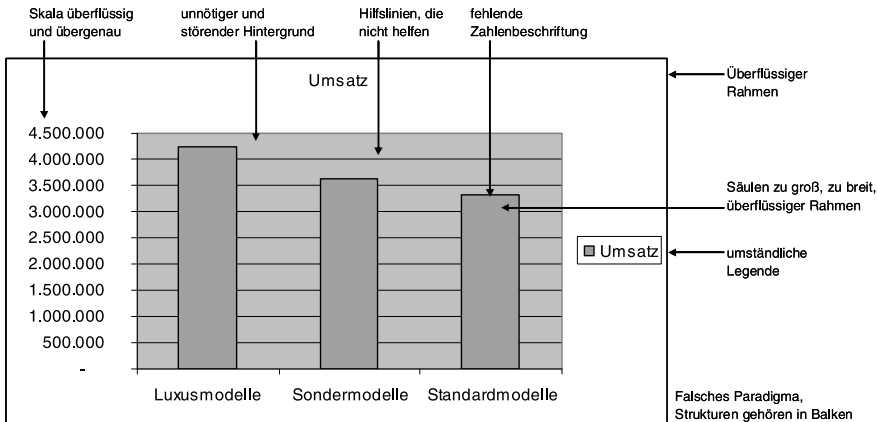


Abbildung 1: Grafischer Schwulst („Chartjunk“): Skala, Raster, Hintergrund, Balkenrahmen, Einfärbung und Legende liefern keinen Beitrag, die eigentlichen Werte muss man erraten

Schon Umsätze von nur drei Produktgruppen wirken nachgerade pathetisch. Wird derlei dünnen Daten obendrein eine ganze Seite in Berichten oder Präsentationen gewidmet, erscheint das unfreiwillig komisch. Ohne Mühe wäre Platz für naheliegende Ergänzungen (z. B. Hintergrundinformationen zur Umsatzentwicklung des Gesamtmarktes) gewesen.

Und mehr noch: Die Grafik ließe sich verlustfrei auf den Satz „Luxusmodelle haben 4,2 Mio. Umsatz erzielt, Sondermodelle 3,6 und Standardmodelle 3,3.“ reduzieren und kann damit ganz entfallen.

2.3 Analysieren heißt Vergleichen

Generell gilt: Zahlen, die wir nicht einordnen können, sind nichtssagend. Wir benötigen einen Vergleichsmaßstab. Wenn wir sonst nichts über das Unternehmen wissen, sind die Umsatzdaten zu den drei Produktgruppen wertlos. Erfahren wir aber, dass wir den Umsatz der Luxusmodelle um 30 % gegenüber dem Vorjahr steigern konnten, dadurch insgesamt auf Platz zwei innerhalb unserer Branche vorgerückt sind, während das Branchenvolumen um 10 % gesunken ist, dann stehen die Daten im notwendigen Kontext und wir können sie bewerten.

Und wir wollen diesen Kontext selbst sehen. Auch wenn daran nichts falsch ist, die Aussage: „Wir verbessern unsere Branchenposition deutlich“ lässt uns unbefriedigt zu-

rück. Kontext erhöht Glaubwürdigkeit. Unsere Empfänger wollen ihre Schlüsse selbst ziehen.

Der Kontext ist selten so spektakulär wie im Beispiel gerade eben. Das muss er auch nicht sein. Den Zusammenhang, das ganze Bild zu sehen statt einen willkürlichen Ausschnitt, das macht die Spannung aus.

Kontext ist vielfältig: Wie sehen die Zahlen im Vergleich zum Vorjahr aus? Ist die Verteilung des Umsatzes zwischen den Bereichen erwartungsgemäß? Gibt es Planwerte dazu? Was bedeuten diese Umsätze für das Ergebnis? Nicht immer steigt der Gewinn, wenn der Umsatz zugenommen hat. Kann ich die Umsätze um Deckungsbeiträge ergänzen? Gibt es Bemerkenswertes auf tieferen Produktebenen?

2.4 Scheingenauigkeit vermeiden

Pathos entsteht auch bei unnötiger Genauigkeit. Sie verstopft ebenfalls den Informationskanal und wirkt oft peinlich unreflektiert. Im Jahr 2004 waren 368.472 der insgesamt 818.271 Todesfälle auf Krankheiten des Kreislaufsystems zurückzuführen. Meint zumindest das Statistische Bundesamt. Die Todesursachenstatistik ist notorisch ungenau. Aufgrund systematischer Mängel wie kleinen Autopsieraten und schlichten Eingabe-, Lese- und Übertragungsfehlern kann sie gar nicht präzise sein. Kritiker halten daher allenfalls die erste Ziffer für wahr.

Da wir Statistiken für umso genauer halten, je mehr Stellen sie uns präsentieren, ist der Übergang zur Manipulation fließend [Bis02a]. Wer exakte Zahlen bringt, suggeriert hohe Glaubwürdigkeit.

Unternehmenszahlen so genau wie möglich festzustellen und zu speichern ist selbstverständlich. Diese Genauigkeit ungeprüft in jeden Bericht zu übernehmen ist schädlich.

	in €	Ist	Plan	Abw.	Abw. %
Brutto-Umsatz		892.645.289	935.339.162	-42.693.873	-4,6%
Rabatte		74.199.130	66.797.643	7.401.486	11,1%
Erlösschmälerung		33.052.701	34.965.234	-1.912.533	-5,5%
Netto-Umsatz		785.393.458	833.576.284	-48.182.826	-10,2%
Lohnkosten		365.852.605	387.825.061	-21.972.456	-5,7%
Materialkosten		16.978.126	16.800.091	178.035	1,1%
Deckungsbeitrag		402.562.728	428.951.133	-26.388.405	-5,6%

Abbildung 2: Millionenbeträge auf den Euro genau tragen wenig zum Erkenntnisprozess bei. Im Gegenteil, sie suggerieren Verlässlichkeit und Glaubwürdigkeit, die daraus abgeleitete Entscheidungen nicht haben können

in Mio. €	Ist	Plan	Abw.	Abw. %
Brutto-Umsatz	893	935	-43	-5%
Rabatte	74	67	7	11%
Erlösschmälerung	33	35	-2	-6%
Netto-Umsatz	785	834	-48	-10%
Lohnkosten	366	388	-22	-6%
Materialkosten	17	17	0	1%
Deckungsbeitrag	403	429	-26	-6%

Abbildung 3: Angemessene Genauigkeit schafft Raum für mehr Daten und Kontext, verbessert die Lesbarkeit und das Verständnis

2.5 Maximale Informationsdichte

Ein erwünschter Nebeneffekt angepasster Genauigkeit ist höhere Informationsdichte. Darin haben viele Managementberichte erheblichen Nachholbedarf. Wie erwähnt: Die Morgenlektüre eines Managers ist häufig dichter als das Berichtswesen seines Unternehmens.

Das ist merkwürdig, denn Dichte ist ein Wert an sich. Überall wird danach gestrebt. Mehr Transistoren auf einer Platine, mehr Pixel je Zoll Bildschirm, mehr Dots per Inch beim Druck, mehr Daten pro DVD, höhere Übertragungsraten in bestehenden Datenleitungen, dichtere Benzin-Luft-Gemische im Motor usw. – Verdichtung ist immer wieder der Schlüssel für höhere Effizienz, Erkennen, Fortschritt.

Hin und wieder wird der geringe Umfang von Informationen z. B. bei Menüstrukturen damit begründet, psychologische Experimente hätten gezeigt, dass wir Schwierigkeiten haben, uns mehr als sieben Dinge auf einmal zu merken. Dabei wird auf ein bekanntes Essay von George Miller [Mil56] verwiesen. Millers Schrift legt jedoch nichts derartiges nahe. Die Experimente bezogen sich auf bedeutungslose Silben, die in keinem Zusammenhang zueinander stehen. Miller empfiehlt im Gegenteil den Kontext von Informationen zu erweitern, um damit unserem Gedächtnis auf die Sprünge zu helfen.

Das Auge-Gehirn-System erfasst Bilder und Muster in rasender Geschwindigkeit. Man spricht von einer Verarbeitungskapazität von 10 Mio. Bits pro Sekunde [Kre06]. Das lässt sich wohl nicht vollständig auf die Verarbeitung abstrakter Daten übertragen. Dennoch sollte nicht ausgerechnet die Darstellung selbst der Flaschenhals in der Informationsverarbeitung sein.

Je dichter wir Informationen auf eine Seite packen, desto weniger müssen wir sie sequenziell anordnen. Das erleichtert nicht nur den Vergleich, sondern schon die Merkfähigkeit.

Dass all dies nichts mit Lesbarkeit zu tun hat, zeigt der Vergleich verschiedener Auflösungen. Eine Schrift in 8 Punkt Größe ist auf einem Bildschirm mit einer üblichen Auflösung von 1.280 x 1.024 Pixeln bereits eine Qual, auf Papier mit 1.200 DPI gedruckt

hingegen gut lesbar. Was lesbar und übersichtlich ist, ist eine Frage der Gestaltung und nicht der Dichte.

Vor diesem Hintergrund ist eine übliche PowerPoint-Präsentation, die auf insgesamt 20 Folien je 5 Stummelsätze versammelt, eine Folter für unser informationshungriges Gehirn. Da wir etwa zehn Mal schneller lesen als sprechen können, lässt sich die Qual nur noch dadurch steigern, dass der Präsentator uns jeden Stummelsatz auch noch vorliest.

2.6 Automation von Analysen, wo immer möglich

Wie aber finden wir in multidimensionalen Datenbanken überhaupt die interessanten Datenkonstellationen, um unsere Fragestellungen zu beantworten? Erst wenn wir diese Informationen gefunden haben, lohnt sich eine anspruchsvolle, informationsdichte Darstellung der Ergebnisse. Die manuelle Suche gestaltet sich naturgemäß oft mühsam und jegliche Automation von Analysen ist prinzipiell willkommen, wenn sie den Suchprozess beschleunigt.

„Automation von Analysen“, damit ist unweigerlich der Begriff des Data Minings verknüpft: Es erhebt den Anspruch, Auffälligkeiten automatisch mit mathematischen Algorithmen in großen Datenbeständen ohne vorab definierte Fragestellung zu identifizieren [BiH96, Bis99]. Im Regelfall werden damit jedoch wieder nur singuläre Erkenntnisse gewonnen und es fehlt die gewünschte Verkettung zu Erkenntnispfaden.

Außerdem bilden die üblichen Data-Mining-Methoden die menschliche Fähigkeit, Zusammenhänge in Daten mit bloßem Auge zu erkennen, nur ungenügend ab. Insbesondere die dichte Platzierung von Informationen innerhalb der Augenspanne befördert eine Vielzahl von intuitiven Analysen durch direkte Vergleiche. So sind Ähnlichkeiten oder Gegenläufigkeiten in zeitlichen Entwicklungen durch das menschliche Auge unmittelbar erkennbar. Diese Fähigkeit zur intuitiven Datenanalyse ist physisch beim Menschen als Ergebnis der Evolution stark ausgeprägt. Eine völlige Automatisierung ist hingegen kaum möglich und nicht effizient.

Dennoch existieren Aufgaben, z. B. das Suchen von initialen Auffälligkeiten in komplexen Datenbeständen, die automatisiert im Sinne eines Data Mining durchgeführt werden können [Bis02b]. Ebenso können mithilfe von Navigationsalgorithmen teilweise automatisch relevante Erkenntnispfade entwickelt werden. Ziel bei der Verwendung von Data-Mining-Methoden muss es daher sein, die Integration solcher Methoden in eine informationsdichte Visualisierung umzusetzen.

3 Wie Berichte berichten

Um zu einem bescheidenen, ernsthaften Berichtswesen zu gelangen, dessen visueller Anspruch im angemessenen Verhältnis zu seinem Steuerungsbeitrag steht, gilt es, sich von tradierten Darstellungsformen zu trennen, den sogenannten Businessgrafiken. Stattdessen plädieren wir für Grafische Tabellen. Sie funktionieren selbst in Kombination mit

Data-Mining-Methoden und eröffnen neue Wege, z. B. im Bereich der Abweichungsanalyse.

3.1 Tod der Businessgrafik

Die gut erkennbaren Zusammenhänge im Managementreporting sind in der Regel sehr schlicht. Im Vordergrund stehen Zahlen, die für Geldbeträge, Mengen, Anteile, Häufigkeiten und Konzentrationen stehen. Eine große Rolle spielen additive Schemata, z. B. in Form einer Bruttogewinn- oder einer Kostenartenrechnung.

Weit verbreitet sind Ergebnisrechnungen, die mithilfe von Standardkosten Deckungsbeiträge auf Basis von Artikeln ausweisen. Komplexität entsteht dann durch eine mitunter hohe Anzahl verschiedener Betrachtungswinkel (Kunde, Region, Produkt etc.).

Insgesamt sind die Bausteine des Managementreporting überschaubar: Ein Analysewert oder ein ganzes Schema lassen sich in verschiedenen Perspektiven betrachten. Eine Perspektive entsteht aus der Kombination von Elementen der vorhandenen Dimensionen, z. B. Kundengruppe X und Artikel Y. Die Perspektiven wiederum können durch Drill Down hierarchisch verfeinert werden. Perspektive und Schema wiederum lassen sich entlang der Zeitachse verschieben, z. B. wenn der Umsatz von Kundengruppe X mit Artikel Y zunächst im Oktober und dann im September betrachtet wird.

Schema, Drill Down und Zeit sind damit die Bausteine, mit denen wir uns in der Hauptsache befassen müssen. Mitunter kommt Raum als weitere Perspektive hinzu (geografische Darstellungen).

Überraschend ist, wie sehr die am weitesten verbreiteten Formen der Darstellung dabei versagen, die für kontextsichere Aussagen unabdingbare Komponente Zeit zu integrieren.

Mit herkömmlichen Mitteln lässt sich die zeitliche Perspektive im Managementreporting nur sehr unbefriedigend abbilden. Nehmen wir an, wir haben Deckungsbeitragsdaten für zwölf Monate im Ist und Plan, mit absoluter und relativer Abweichung. In tabellarischer Form füllen diese Daten vier Seiten DIN A4.

	Ist	Plan	Abw.	Abw. %
Brutto-Umsatz	15.385	16.411	-1.026	-6,3%
Rabatte	1.003	1.056	-53	-5,0%
Erlösschmälerung	537	595	-58	-9,7%
Netto-Umsatz	13.844	14.759	-915	-6,2%
Lohnkosten	6.756	7.228	-472	-6,5%
Materialkosten	292	309	-17	-5,6%
Deckungsbeitrag	6.796	7.223	-427	-5,9%

Abbildung 4: Ein typisches Schema im Reporting



Abbildung 5: Mit einer normalen Tabelle nicht zu schaffen:
Deckungsbeitragsdaten für 12 Monate mit Ist, Plan, absoluter und relativer Abweichung

Das ist nicht zu handhaben. Nur Merkakrobaten wird es gelingen, im zeitlichen Vergleich Schwankungen, Höhen und Tiefen zu erfassen. Kurzum, wir sehen den Wald vor lauter Bäumen nicht. Aber auch mit einer Businessgrafik kommen wir nicht weit:

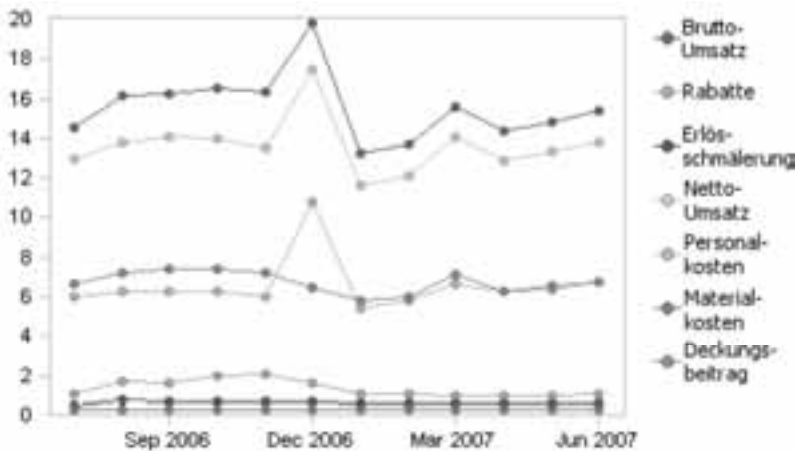


Abbildung 6: Auch eine Businessgrafik stößt schnell an ihre Grenzen:
Beschriftung und Skalierung werden nur unbefriedigend gelöst

Obwohl wir nur die Ist-Spalte unserer Daten abbilden, scheitern wir. Wir müssen aufwändige Dechiffrierarbeit leisten, nur um die Bedeutung jeder einzelnen Linie zu erkennen. Das Auge springt hin und her. Manchen wird es schwer fallen, die Farbunterschiede rasch zuzuordnen. Auch die Skalierung ist unbefriedigend. Der Verlauf von Posten mit kleinerem Wertebereich bleibt unklar und das additive Deckungsbeitragschema ist nicht mehr erkennbar.

3.2 Tufte's duftete Idee: Wortgrafiken (Sparklines)

Wortgrafiken oder Sparklines, wie ihr Erfinder Edward Tufte sie nennt, lösen das Problem der Integration der zeitlichen Perspektive auf elegante Weise. In Wortgröße integrieren sie die Zeit in die Tabelle. Der daneben stehende Zahlenwert repräsentiert den Wert des letzten Balkens in der Reihe. Weicht dieser Wert nicht zu dramatisch von den übrigen ab, so genügt das für eine ausreichende Quantifizierung der Reihe insgesamt [Bis05].

		Ist	Plan		Abw.	Abw. %
Brutto-Umsatz		15.385	16.411		-1.026	-6,3%
Rabatte		1.003	1.056		-53	-5,0%
Erlösschmälerung		537	595		-58	-9,7%
Netto-Umsatz		13.844	14.759		-915	-6,2%
Lohnkosten		6.756	7.228		-472	-6,5%
Materialkosten		292	309		-17	-5,6%
Deckungsbeitrag		6.796	7.223		-427	-5,9%

Abbildung 7: Sparklines integrieren die Zeit in tabellarische Darstellungen

Es ist erstaunlich, dass Sparklines die Lesbarkeit von Zeitreihen durch die Miniaturisierung verbessern und nicht verschlechtern, wie der Vergleich mit den Alternativen oben zeigt. In der rein tabellarischen Darstellung ist das Muster des Jahresverlaufs gar nicht erst erkennbar, sie kann allenfalls der buchhalterischen Überprüfung auf Richtigkeit der Werte dienen. Die Businessgrafik wiederum ist extrem schlecht ablesbar.

Die Größe von Sparklines erlaubt es zudem, sie in Sätze zu integrieren: „Der Jahresumsatz der Deutschen Telekom seit 1993 bis 2007 mit zuletzt 62,5 Mrd. Euro ...“. Sofort wird klar, dass der Telko-Umsatz einige Jahre auf ähnlichem Niveau blieb, bevor er stetig anstieg. Ohne Sparkline hingegen wäre das so dicht nicht zu vermitteln.

Damit Sparklines funktionieren, gilt es ein paar Regeln einzuhalten. Erst die geschickte Anordnung von Zeitraum, letztem Wert und der Grafik ergeben die Sparkline. Gut platziert zeigen sie das Auf und Ab des amerikanischen Haushaltsdefizits von 1983 bis zu den -354,2 Mrd. Dollar im Jahr 2007, die Entwicklung des Euro gegenüber dem Dollar seit seiner Einführung als Buchgeld an Neujahr 1999 bis Dezember 2007 1,47 mit seinem katastrophalen Einbruch nach der Einführung als Bargeld am 01.01.2002 um fast 25 % oder die gewonnenen (Balken nach oben) und verlorenen Spiele (Balken nach unten) des deutschen Fußballmeisters FC Bayern München in der Saison 2007/2008.

3.3 Nacharbeitungsfreies Reporting mit Grafischen Tabellen

Auf die Vorteile von Grafiken wollen wir nicht verzichten. Gut gemacht, lenken sie das Auge und transportieren Größenverhältnisse und Auffälligkeiten schneller als viele Tabellen. Das gelingt im Wesentlichen dadurch, dass Wertverhältnisse proportional in Längenverhältnisse übersetzt werden. Die Leistung von Säulenreihen für die Darstellung von Zeitreihen haben wir uns mit Sparklines bereits erschlossen (gleiches gilt für Sparklines auf Basis miniaturisierter Liniendiagramme).

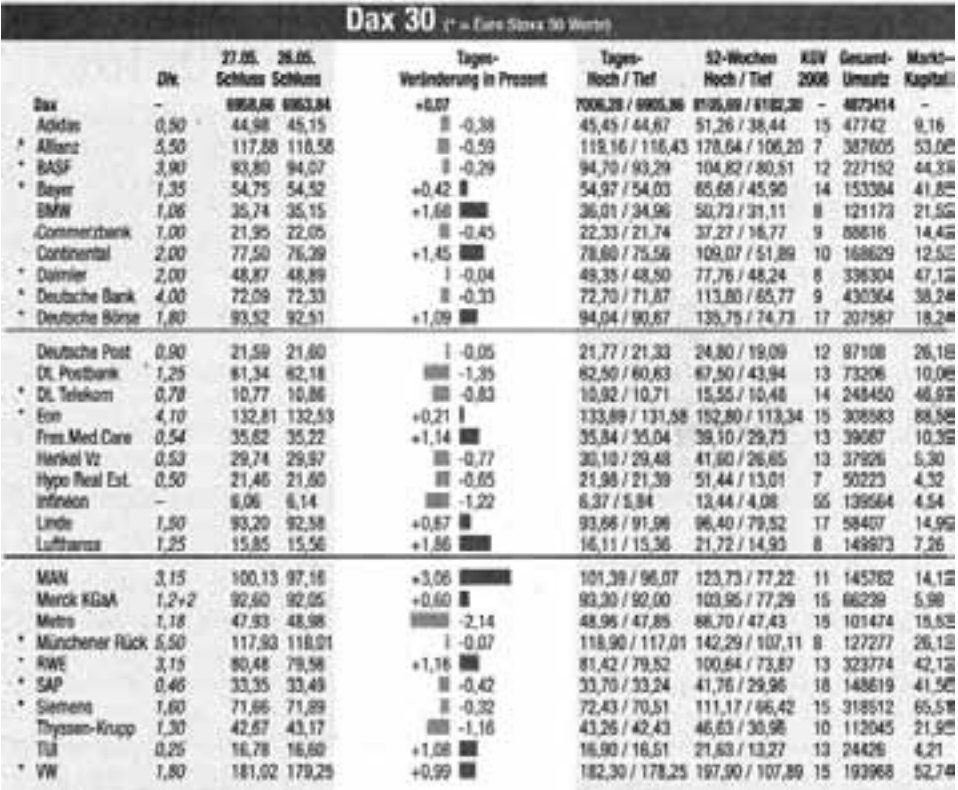


Abbildung 8: In vielen Tageszeitungen bereits eine Selbstverständlichkeit: kompakte Integration von Text, Zahl und Balkendiagramm (Quelle: Die Süddeutsche Zeitung, 28.05.2008, Ausschnitt)

Die meisten (stehenden) Säulen der typischen Businessgrafiken kann man durch (liegende) Balken ersetzen. Balken müssen nicht höher als die Schrift und nicht breiter als eine normale Zelle mit Zahlen sein. Damit sind die Voraussetzungen erfüllt, Text, Zahl und Grafik kompakt und informationsdicht in eine Tabelle zu integrieren.

Damit sind wir beim Konzept der grafischen Tabellen. Sie integrieren die aus Diagrammen bekannten Zeichnungselemente in der Regel in verkleinerter Form so in die Tabellenzellen, dass sich die Beschriftung durch die bereits vorhandenen Zahlenwerte von selbst ergibt. Im Ergebnis werden so die Vorteile von Tabellen und Grafiken verknüpft und gleichzeitig ihre Nachteile vermieden.

Wie das folgende Beispiel zeigt, wird die hohe Informationsdichte von Tabellen durch Sparklines nochmals dramatisch erhöht, durch die Integration von Balkenelementen aber kaum verringert. Umgekehrt entstehen grafische Darstellungen mit enormer Informationsdichte. Diese Tabelle wuchs in der Fläche um den Faktor 1,5 gegenüber der Ursprungstabelle ohne grafische Elemente und repräsentiert jetzt jedoch 6,5-mal mehr Werte.

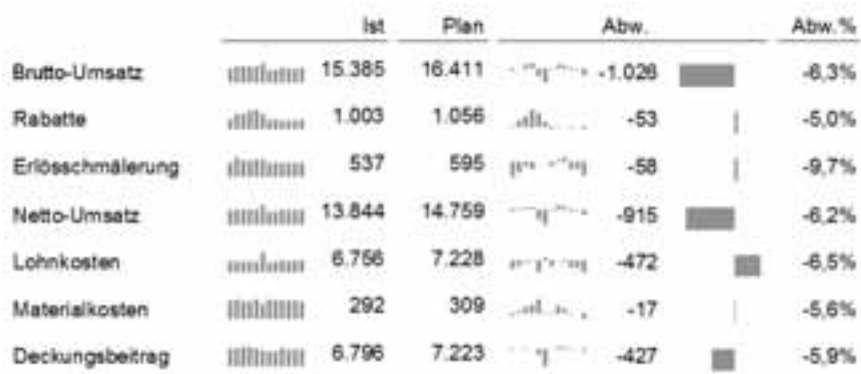


Abbildung 9: Grafische Tabelle mit Sparklines und integrierter Balkengrafik

Grafische Tabellen sind in hohem Maße automatisierbar. In Grafiken hängt es von den Daten ab, ob Legenden und Beschriftungen funktionieren, daher müssen sie bei jedem Berichtslauf nachgearbeitet werden. Bei grafischen Tabellen sind die Spaltenbreiten ohne Schaden fixierbar. Die grafischen Elemente passen sich dem an.

Grafische Tabellen erhöhen die Wirtschaftlichkeit des Managementreporting durch Berichtstypen, die keiner mühsamen Nacharbeit mehr bedürfen. Kosten werden dadurch reduziert, dass auf einen Teilschritt ganz verzichtet werden kann. In der anschließenden Berichtsverteilung bieten sich weitere Ansatzpunkte. Häufig entstehen Berichte durch Variation eines Grundtyps, der an die jeweiligen Empfänger angepasst wird. Diese iterativen Prozesse können an den Rechner delegiert werden, da sie einer einfachen Zuordnungslogik folgen. Wir substituieren in diesem Fall teure menschliche Arbeitszeit durch billige Rechenzeit.

3.4 Data Mining im Reporting

Herkömmliche Tabellen zeigen den Status Quo eines Datenausschnitts. Wie wir zeigen konnten, sind Vergleiche nur in engen Grenzen möglich. Sparklines erweitern diese Grenzen dramatisch. Sie integrieren die wichtige Dimension „Zeit“. In der Software DeltaMaster sind grafische Tabellen zudem mit Data-Mining-Verfahren für die automatische Navigation verknüpft. Ziel ist, selbsttätig neue Erkenntnispfade zu entdecken, die in den Datenkontext eingebettet bleiben.

Das folgende Beispiel illustriert die Vorgehensweise. Der Benutzer startet mit einer Pivottabelle, die beispielsweise eine Deckungsbeitragsrechnung mit Plan-Ist-Abweichungen darstellt. Integrierte Sparklines zeigen den historischen Verlauf der Kennzahlen und liefern den nötigen Kontext zu den aktuellen Daten. Im Beispiel ist eine Zunahme bei den Rabatten in den letzten Perioden erkennbar.

Zusätzlich werden die Daten mit robusten statistischen Heuristiken daraufhin überprüft, ob Trends vermutet werden können. Ist das der Fall, so wird ein Pfeilsymbol mit einer der Trendrichtung angepassten Steigung dargestellt. Liegt kein Trend vor, wird dies entsprechend vermerkt („kT“). Optional visualisieren Balkendiagramme die Größenverhältnisse ausgewählter Kennzahlen. Weitere Visualisierungsformen, wie z. B. Füllstandsanzeigen, ergänzen das Methodenportfolio (hier nicht abgebildet).

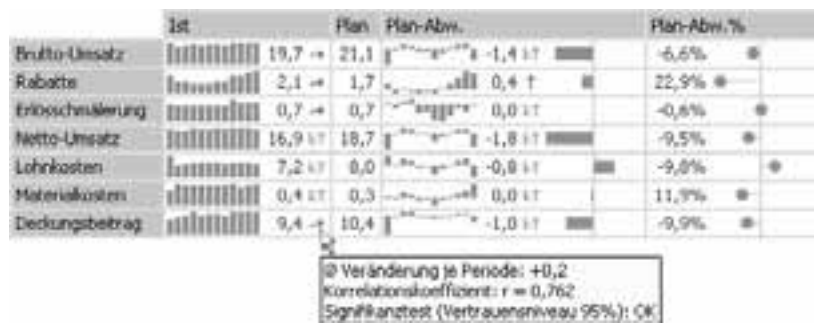


Abbildung 10: Grafische Tabelle in DeltaMaster



Abbildung 11: Auswahl des Data-Mining-Verfahrens „Automatische Navigation“

Zur detaillierten Untersuchung der Abweichungen des Ist-Deckungsbeitrags (DB) vom Plan-DB wird die automatische Navigation per Maus aufgerufen. Dabei untersucht ein Data-Mining-Algorithmus die Relevanz aller im Datenmodell verfügbaren Einflussfaktoren auf den Deckungsbeitrag. So stehen z. B. Produkte, Kunden, Stoffgruppen und Vertretergruppen im Datenmodell als potenzielle Erklärungsdimensionen zur Verfügung.

DeltaMaster identifiziert automatisch den Faktor, der den größten Erklärungsbeitrag für die Kennzahl bietet. Eine Analysedimension gilt als umso auffälliger, je ungleichmäßiger der untersuchte Wert auf die Elemente der Dimension verteilt ist. Das Objekt mit

dem heterogensten Bild wird von der automatischen Navigation ausgewählt. Die Pivot-tabelle wird partiell erweitert (vgl. folgende Abbildung), um den Einflussfaktor darzu-stellen. Im Beispiel erklärt die Modelldimension „Kunden“ die Abweichung des De-ckungsbeitrags am stärksten. Es zeigt sich, dass die Planabweichung im Süden Europas größer ist als die kumulierte Abweichung für Gesamteuropa. Die anderen Regionen weisen kompensierende Effekte mit positiver DB-Abweichung auf.

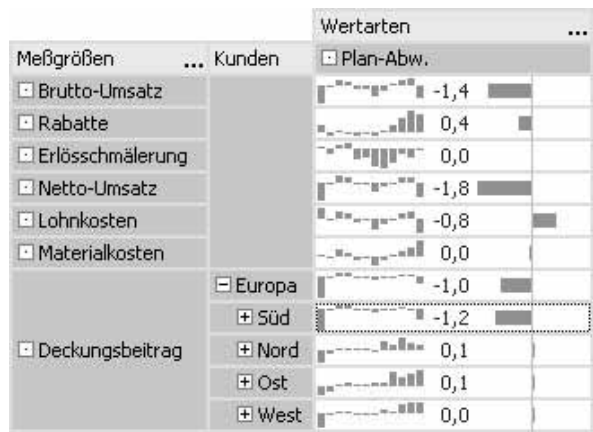


Abbildung 12: Partiiell erweiterte Pivot-tabelle als Ergebnis der automatischen Navigation

Die Folgefrage: „Welcher Einflussfaktor verursacht in einer Region (hier: Südeuropa) maßgeblich die Deckungsbeitragsabweichung?“ beantwortet der nächste automatische Navigationsschritt. Das folgende Beispiel zeigt drei aufeinanderfolgende Schritte. Er-gebnis ist ein Erkenntnispfad, an dessen Ende die Hauptverursacher für die Deckungs-beitragsabweichung zu finden sind. Im Beispiel verursacht der Markt Südeuropa mit Sondermodellen in Chrom den größten Teil der Deckungsbeitragsabweichung in der aktuellen Periode.

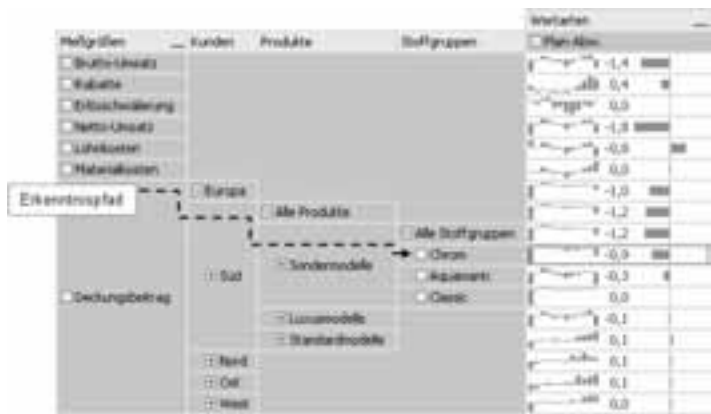


Abbildung 13: Mehrstufiger Erkenntnispfad als Ergebnis der automatischen Navigation

Zu jedem Zeitpunkt kann der Nutzer vom empfohlenen Navigationspfad zugunsten einer anderen Fragestellung abweichen und einem anderen Erkenntnispfad folgen. Das sichert die nötige Flexibilität.

4 Fallstudie Transportkostenmanagement

Das nachfolgende Fallbeispiel ist dem Controlling eines Logistikdienstleisters entlehnt. Im Mittelpunkt stehen Transportleistung und -kosten. Charakteristisch ist die Betrachtung nach mehreren Perspektiven, wie z. B. Produkte, Verlager, Transportarten und Verkehrsträger. Kennzahlen sind Transportmengen, Distanzen sowie Erlös- und Kostenarten für die Transportaufträge. Daraus lassen sich mehrstufige Deckungsbeitragsrechnungen aufbauen. Ziel ist, die besonders profitablen Kunden, Strecken und Produkte zu identifizieren.

Die sogenannten KPI (Key Performance Indicators) werden mit informationsdichten Übersichten überwacht. Abbildung 14 zeigt ein Beispiel zu den Finanz- und Leistungsindikatoren. Die Möglichkeiten von Grafischen Tabellen, Trenderkennung und Sparklines werden konsequent genutzt.



Abbildung 14: Beispiel eines Cockpits für die Finanz- und Leistungsperspektive eines Logistikdienstleisters

In einer anderen KPI-Auswertung (vgl. Abbildung 15) steigt der Indikator „Leistungskosten/Transportmittel“ von einer Periode zur nächsten für alle Häfen deutlich an, wie der Sparkline zu entnehmen ist.

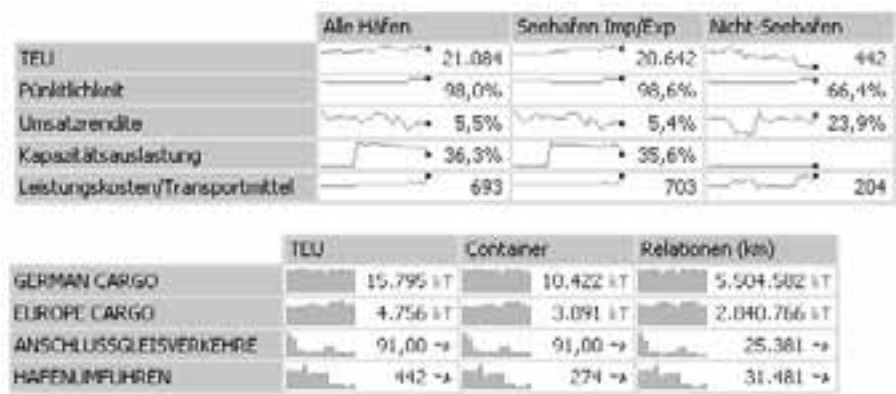


Abbildung 15: Ausschnitt einer KPI-Auswertung

Ein erster Blick auf die zeitliche Entwicklung dieses Schlüsselindikators (vgl. Abbildung 16) ergibt, dass die Kosten weitgehend konstant blieben, jedoch die Menge an Transportmitteln stark gesunken ist. Daher steigt der Indikator entsprechend stark an.

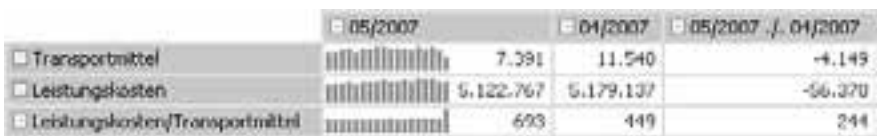


Abbildung 16: Pivottabelle zur Analyse eines KPI

Daraus ergeben sich zwei Fragen: Warum sind die Transportmittelmengen so stark gefallen und lassen sich Verursacher feststellen? A priori besteht keine Vermutung, wo der Verursacher zu suchen ist. Die automatische Navigation wird daher ohne Vorgabe durch den Benutzer gestartet. Abbildung 17 zeigt das Ergebnis. Über vier Drill-Down-Stufen findet der Nutzer rasch den Grund für die Abweichung: Die Veränderung der Transportmittelmenge von insgesamt mehr als 4.100 Einheiten wurde zu fast 25 % (rund 1.000 Einheiten) von einer einzigen Kundengruppe, „International Container“, verursacht (vgl. Pivottabelle). Ein Gespräch mit dem für den Kunden zuständigen Salesmanager schließt sich an, um möglichen Gründen nachzugehen, die sich nicht in Zahlen fassen lassen.

Weiterhin zeigt sich, dass rund 75 % der verursachenden Transporte offensichtlich innerdeutsche Transporte waren. Empfangsland (Drill-Stufe 1) und Versandland (Drill-Stufe 2) ist jeweils Deutschland. Auf diese innerdeutschen Transporte entfallen rund 3300 Einheiten.

6 Literaturverzeichnis

- [BiH96] Bissantz, N., Hagedorn, J.: Data Mining im Controlling (Dissertationen), Arbeitsberichte des Instituts für mathematische Maschinen und Datenverarbeitung (Informatik), Band 29, Nr. 7, Erlangen 1996.
- [Bis99] Bissantz, N., Aktive Managementinformation und Data Mining: Neuere Methoden und Ansätze In: (Chamoni, P. et al., Hrsg.), Analytische Informationssysteme: Data Warehouse – OLAP – Data Mining. 2. Auflage, Berlin u. a. 1999: S. 375-392.
- [Bis00] Bissantz, N.: Praxiserfahrungen mit Data-Mining-Projekten. In: Grothe, M., Gentsch, P., Business Intelligence – Aus Informationen Wettbewerbsvorteile gewinnen, München et al. 2000, S. 188-206.
- [Bis02a] Bissantz, N.: So lügt man mit Grafiken, Arbeitspapier Bissantz & Company GmbH, Nürnberg 2002, http://www.bissantz.de/pub/Luegen_mit_Statistiken.pdf
- [Bis02b] Bissantz, N.: Stand und Weiterentwicklung softwaregestützter Datenanalysen im betriebswirtschaftlichen Umfeld. In: (Hannig, U., Hrsg.), Knowledge Management und Business Intelligence. Berlin u. a. 2002, S. 229-245.
- [Bis05] Bissantz, N.: Sparklines: Wie grafische Substantive und Verben das Controlling vereinfachen und beschleunigen. Controller Magazin 30, 2005, S. 484 u. 502.
- [Bis06] Bissantz, N.: Sind Fußballfans klüger als Manager?, <http://blog.bissantz.de/basketball>, 13.11.2006.
- [Few06] Few, S.: Information Dashboard Design, Sebastopol 2006.
- [Hic07] Hichert, R.: Managementberichte verständlich gestalten – Empfehlungen für Verbesserungen, CFO aktuell, August 2007, S. 173-175.
- [Hic08] Hichert, R.: Klare Botschaften und verständliche Darstellungen verschaffen dem Controller Gehör, Zeitschrift für Controlling & Management – ZfCM 1/2008, S. 15-17.
- [Kre06] Kreeger, K.: Penn researchers calculate how much the eye tells the brain, http://www.eurekalert.org/pub_releases/2006-07/uops-prc072606.php
- [MeG02] Mertens, P., Griesse, J.: Integrierte Informationsverarbeitung 2, Wiesbaden, 2002.
- [Mil56] Miller, G. A.: The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, Psychological Review 63, 1956, S. 81-97.
- [Tuf06] Tufte, E. : Beautiful Evidence, Cheshire 2006.

Organising Accountabilities for Data Quality Management – A Data Governance Case Study

Kristin Weber*, Lai Kuan Cheong**, Boris Otto*, Vanessa Chang**

*Institute of Information Management
University of St. Gallen
Müller-Friedberg-Strasse 8
9000 St. Gallen, Switzerland
{kristin.weber | boris.otto}@unsig.ch

**School of Information Systems, Curtin Business School
Curtin University of Technology
Kent Street, Bentley,
Western Australia 6102, Australia
lkcheong@yahoo.com
vanessa.chang@curtin.edu.au

Abstract: Enterprises need corporate data quality management (DQM) that combines business-driven and technical perspectives to respond to strategic and operational challenges that demand high-quality corporate data. Hitherto, companies have assigned accountabilities for DQM mostly to IT departments. They have thereby ignored the organisational issues that are critical to the success of DQM. With data governance, however, companies implement corporate-wide accountabilities for DQM that encompass professionals from business and IT. This study examines a large organisation that has adopted an ad-hoc data governance model to manage its data. It was found that its DQM efforts were hampered mainly by the lack of clear roles and responsibilities and the lack of mandate to carry out data quality improvement initiatives. In order to promote effective DQM, this research identifies a data governance structure with the emphasis on collaboration between business and IT to support organisations.

1 Introduction

Companies are forced to continuously adapt their business models. Global presence requires harmonised business processes across different continents, customers ask for individualised products, and service offerings must be industrialised [cf. KÖ06]. These factors certainly impact the business process architecture and the IT strategy of organisations. Ultimately, however, data of high quality is a prerequisite for fulfilling those changing business requirements and for achieving enterprise agility objectives [NL06a]. In addition to such strategic factors, some operational domains directly rely on high-quality corporate data, such as business networking [Ve00; Ma04; Te04], customer man-

agement [ZG03; RC05; CM06], decision-making and business intelligence [SZW03; PS05], and regulatory compliance [Fr06].

Data quality management (DQM) focuses on the collection, organisation, storage, processing, and presentation of high-quality data. In addition, there are organisational issues that must be addressed, such as maintaining sponsorship, managing change and expectation, defining accountabilities, enforcing mandates, and handling political issues [Wa98a; En99; Ep06; Th06]. Despite those organisational aspects, responsibility for improving data quality and managing corporate data is often assigned to IT departments [Fr06]. Also, many companies try to cope with data quality (DQ) issues by simply implementing data management or data warehouse systems. Surveys on data warehousing failures reveal that organisational rather than technical issues are more critical to their success [WFA04].

Successful DQ programs identify the organisational processes behind DQ [BN07]. In order to address both organisational and IT perspectives an integrated approach to DQM is required. With data governance, companies implement corporate-wide accountabilities for DQM that encompass professionals from both business and IT. Data governance defines roles and assigns accountabilities for decision areas to these roles. It establishes organisation-wide guidelines and standards for DQM and assures compliance with corporate strategy and laws governing data.

There are only few academic resources about data governance. Apart from a few DQM approaches dealing with accountabilities [Re96; En99], an elaborate analysis of the interaction of roles and responsibilities, and the design of decision-making structures is missing. For our research, we therefore incorporate data governance sources from consultants, analysts and practitioners [e.g., Sw05; De06a; DL06a; MS06; NL06b; Ru06; Th06; BN07].

The conclusion from these sources is that data can be managed more effectively and successfully through the adoption of a data governance structure. This paper outlines a case study of an Australian utility company (“Water Inc.”)¹ to determine the justification for formal data governance. Our main contribution is to demonstrate how formal data governance structures contribute to successful DQM based on this case study. The paper exposes how the state of the art in data governance was adopted by the company in order to build a virtual, boundary-spanning organisation for managing data quality effectively.

The remainder of the paper is structured as follows: Section 2 introduces related work on data quality management and data governance. It provides an overview of formal data governance structures. Section 3 outlines the research methodology and the case study company “Water Inc.”. Section 4 describes the initial situation at “Water Inc.” and details its data governance approach. Section 5 discusses the insights of the case study. The last section, Section 6, summarises the paper and its contribution.

¹ Name of the company changed by the authors.

2 Background

2.1 Data Quality Management and Data Governance

We refer to data quality management as quality-oriented data management, i.e., data management focusing on collecting, organising, storing, processing, and presenting high-quality data². The importance of organisational issues, such as maintaining sponsorship, managing change and expectation, defining accountabilities, enforcing mandates, and handling political issues, result from conflicts between different departments, lines of business, and legal entities in which DQM operates. DQM is boundary-spanning and provides many stakeholders (e.g., CxOs, sales, controlling, procurement, IT, business units, customers, public authorities) with high-quality corporate data by balancing their different interests (e.g., company-wide requirements, laws and regulations vs. local and regional differences). Because of these particularities of DQM, large multi-business companies are likely to have difficulties with institutionalising DQM, i.e., defining accountabilities, assigning people accountable for DQM within the organisational structure, and enforcing DQM mandates throughout the company.

Data governance – as part of DQM – addresses these particular issues within corporate structures. Data governance specifies the framework for decision rights and accountabilities to encourage desirable behaviour in the use of data³. Hence, data governance encompasses two aspects: first, establishing accountabilities for DQM; and second, defining corporate-wide guidelines and standards for DQM. This paper focuses on the first aspect, because this aspect of data governance causes difficulties in most companies. Little practical guidance and few case studies of successful implementations exist. Furthermore, established accountabilities are the pre-requisite for the second data governance aspect – defining and implementing guidelines and standards.

Academic research on data governance is in its infancy. DQM approaches, such as Total Data Quality Management (TDQM) [Wa98b; Wa98a; HLW99], insufficiently address the allocation of accountabilities. They mainly outline DQM activities and decision areas. TDQM defines only one accountable role, the information product manager, which ensures that relevant, high-quality information products are delivered to information consumers. Few DQM approaches, such as Total Quality data Management (TQdM) by [En99] and DQ for the information age by Redman [Re96] deal with more than one accountable organisational position or role, roles related to several organisational levels, and their tasks and responsibilities.

However, an elaborate analysis of the interaction of roles and responsibilities, and the design of decision-making structures is missing. Hence, companies might find it difficult

² The term data is often distinguished from information by referring to data as “raw” or simple facts and to information as data put in a context or data that has been processed [HLW99; PS05]. In line with most data or information quality publications, we use the terms data and information interchangeably throughout the paper.

³ In the absence of academic definitions of data governance, this definition was adapted from the IT governance definition of [We04].

to establish and maintain organisational structures designed to assure and sustain high-quality data throughout the enterprise. Findings of a recent survey among data management professionals indicate that data governance is rarely adopted [Ru06]. Only 8% of respondents had deployed a data governance initiative, 17% were in the design or implementation phase.

2.2 Organising Data Quality Management

Hierarchical organisational structures fail to support corporate DQ in situations where data is used across organisational boundaries [Th06, 25], such as ensuring regulatory compliance or supporting global process harmonisation. Therefore, data governance establishes a “virtual organisation”. It defines roles⁴ and their responsibilities for DQM across organisational boundaries. It establishes committees⁵ to make important unbiased DQM-related decisions to achieve the best result for the whole organisation.

For identifying the roles and organisational bodies involved in DQM, Thomas [Th06, 81ff] distinguishes three kinds of data governance approaches. In *Governance via Management* managers execute decision-rights and set data-related rules. The governance organisation is the existing organisational structure. Responsibilities for DQM are not formalised. *Governance via Stewardship* is a more formal approach that provides separate roles and responsibilities. The governance organisation is made up of a hierarchy of data stewards and data owners. Data stewards typically define, create, or use data and may formulate rules. The *Governance via Governance* approach clearly distinguishes between governors (make rules and resolve issues) and data stewards (work with data, ensure compliance with rules, and raise issues). Governance bodies complement hierarchical management structures to address boundary-spanning DQM issues.

Data governance surveys [Ru06], case studies [De06b; DL06b], and reports and books from analysts and consultants [Sw05; DL06a; MS06; NL06b] follow that governance via governance approach. These sources distinguish between three and five organisational roles building the data governance organisation. Dyché and Levy [DL06a] and English [En99] describe more specialised roles – they distinguish twelve and nineteen roles respectively.

The analysis of these sources results in a set of four typical roles and one committee – the data quality board. Table 2 describes the roles and the committee. It provides a short description, the level and part of the organisation to which the roles typically belong, and the alternative names found in the sources. Names in brackets only partly match with either the description or organisational assignment.

⁴ A *role* bundles different business tasks that are carried out by a single person (employee) or an organisational unit as well as their area of responsibility and competencies.

⁵ A *committee* is a politically-neutral organisational body that unites stakeholders from different parts of the organisation.

Role	Description	Organisational Assignment	Sources
Executive Sponsor	Provides sponsorship, strategic direction, funding, advocacy and oversight for DQM	Executive or senior manager (e.g., CEO, CFO, CIO)	Strategic information steward [En99], executive level [NL06b], executive sponsor [MS06], (executive council) [De06b]
Data Quality Board	Decides for corporate-wide standards and controls its implementation	Committee, chaired by chief steward, members are business unit and IT leaders as well as data stewards	Business information stewardship team [En99], data governance council [DL06a; MS06], data governance committee [Ru06], GRCS board [De06b], trustee council [DL06b], (legislative level) [NL06b]
Chief Steward	Puts the board's decisions into practice, enforces the adoption of standards, helps establish DQ metrics and targets	Senior manager with data management background	Master data coordinator [Sw05], director of data management [DL06a], chief steward [MS06], corporate steward [Ru06], lead stewards [De06b], (data czar) [DL06a]
Business Data Steward	Details the corporate-wide DQ standards and policies for his or her area of responsibility from a business perspective	Professional from business unit or functional department	Information professionals [Re96], business information steward [En99], business data steward [DL06a], subject area steward [NL06b], master data lead [Sw05], domain steward [Ru06], business steward [MS06], subject matter expert [DL06b]
Technical Data Steward	Provides standardised data element definitions and formats, profiles source system details and data flows between systems	Professional from IT department	Database steward & information architecture steward [En99], technical steward [MS06], source system data steward [DL06a]

Table 2: Set of Data Quality Roles

The actual number of roles may vary from company to company. However, the roles presented build a balanced and useful set when focusing on the strategic notion of corporate DQM. The data governance organisation is virtual, i.e., it does not create a new organisational structure; it aligns and coordinates existing roles and cooperates with established committees and roles, such as IT governance boards, central purchasing committees, data architects or data security officers.

Against this background, this paper describes a case study of an Australian utility company. Several incidents that took place related to the lack of maintenance on network asset were the main business driver for starting a DQM initiative. Data for network assets need to be of high quality for the maintenance program to be planned efficiently. This paper outlines the process of implementing a data governance organisation within the company and how it benefited from a formal data governance (via governance) approach in its DQM initiative.

3 Research Methodology

3.1 Case Study Research

We adopted a case study research method which is particularly suitable for understanding phenomena within their organisational context [BGM87; Ei89; Yi02]. Case study research allows researchers to carry out their studies in a natural setting, learn about the actual process of managing data quality and generate theories from practice. This also allows researchers to answer the ‘how’ and the ‘what’ questions in order to understand the nature and complexity of the processes taking place. This is an appropriate research method for an area that had few previous researches [BGM87]. As outlined above, this is particularly true for data governance. Furthermore, data governance is an emerging topic that needs further development and research [WO07].

The chosen utility company (“Water Inc.”) manages data from various disparate systems or sources, and the data involves different line of business units from multiple users and stakeholders. The current user base is more than 100 users and data quality is the priority concern for the organisation. The data collected from interviews with the IT Managers and Data Managers were primarily qualitative in nature. The interviews were recorded and later transcribed and analysed. The interview questions were designed to ascertain the IT view of data-related issues and problems and the business view of data-related issues and problems within the organisation. In addition, the study also investigated the resolution and mediation methods that the organisation employed; and the deployment of such methods. Additional information was also gathered from reviewing data management documentation of the organisation. The documentation review provided insight into the style of data quality management, methodology used for developing and enhancing application systems, and the interaction between IT and business.

3.2 Water Inc.

Water Inc. is a large utility organisation in Australia with complex data integration issues. This organisation provides essential services to approximately 890,000 industrial, commercial, and residential customers. Its main mission is to provide reliable supply by maintaining the network and to restore the supply in the event when the supply was interrupted. It is also responsible for building new network to meet the demands of existing and future customers.

It employs more than 1,850 core staff with an asset base of nearly AU\$ 3 billion. It operates in a regulated market to ensure quality of service and the accessibility of its network to all service retailer and supplier. Recently, Water Inc. was split into four different business entities (regional and metropolitan retail business, generation business and network/infrastructure (asset) management). Through this restructure, it had been charged with AU\$ 2.23 million of investment to increase the network reliability by 25% over the next four years.

Water Inc. is obliged to regularly report to a government regulatory authority. The government regulatory authority requires maintenance plans and proof that the company had

performed its duties in maintaining public safety, ensuring reliable supply, and efficient management of its network infrastructure. The building and maintenance of the network infrastructure requires the cooperation of several divisions: the asset management division (setting the strategy for managing network assets), field services (packaging of inspection and maintenance program), and works delivery (planning and construction of new infrastructure).

A few years ago, several incidents related to the lack of maintenance on network asset affected public safety and reliability of supply. This resulted in an inspection of the conditions of the network assets. Data for these assets need to be current and accurate so that a maintenance program can be planned efficiently. With this maintenance program, the condition of the assets can be accurately determined and aging assets can be targeted for replacement ensuring the reliability of the service network.

Water Inc. is also in the process of migrating its asset related data from its legacy asset management systems to Ellipse, an Enterprise Resources Planning (ERP) system and is also planning to replace its custom-built Geographical Information System (GIS) with an off-the-shelf GIS. The successful migration of these legacy systems depends largely upon the quality of the asset and spatial data. For example, in order to locate assets the XY coordinates and GPS references must be accurate; in order to identify the number of customers affected by an asset malfunction the connectivity of the network must be accurate. Recently, a dedicated Data Management department was created to manage asset data and its related information systems for the Asset Management division. This department is responsible for ensuring data quality and providing strategic direction for the asset related data initiatives. It consists of three teams: First, the Data Management (DM) team is responsible for providing strategic and tactical direction on data related issues, conducting quality assurance and data cleansing activities; second, the Data Services team enters data for asset related application systems (distribution network) and provides underground asset information to external parties conducting trenching activities; and third, the GIS Strategist team is responsible for the strategic direction of any spatial data related issues, such as positional accuracy and graphical representation of data.

4 Data Governance at Water Inc.

This research focuses on the activities within the DM team at Water Inc. This team acts as data stewards. It has knowledge of the business processes and an in depth knowledge of the asset related information systems. Currently, it is facing difficulties with managing asset related data as an enterprise asset. The DM team feels it is managing data in a reactive and ad hoc manner, it has no direct access to its source database, it has difficulties obtaining consensus on data related issues, its data improvement projects were mostly overrun and over budget, and any attempt to set data-related standards had not been taken seriously by other groups of the business and IT. In light of this, the DM team started a data governance initiative and the following remediations were instigated to address these issues:

- *Reactive data management.* The DM team establishes a data quality strategy to improve and maintain data quality. Based on the strategy, a work plan for IT related project and resource management is set-up.
- *Asset data not easily accessible.* The asset data is stored in several disparate databases and maintained by different application systems.
- *Data improvement project mostly overrun and over budget.* A dedicated project manager is appointed to liaise between the business and IT and to ensure that the data stakeholders' expectation and communication are managed effectively. There is still a lack of IT tools for data profiling and data cleansing.
- *Difficulty in setting data standards.* Data standards are important for the migration of data from legacy systems to the new ERP system [CS01]. The DM team plans to purchase a metadata repository.
- *Difficulty in obtaining consensus on data related issues.* Data users are spread across the organisation, spanning different divisions. A data custodianship policy was drawn up but was difficult to put into action as it did not have a governance structure to enforce this policy.

This data governance model corresponds to the governance via stewardship approach [Th06]. However, this form of governance is not working for the DM team due to the following reasons: first, a lack of mandate from senior executives results in no power to act; second, sometimes DQM projects have no priority as senior executives do not understand their importance; and third, a lack of clear roles and responsibilities. Obviously, there is no connection between those at the operational level who knows the problem and those who have the power to make decisions but are removed from the problem. In order to address the issue, the DM team put together a plan for the governance via governance approach of data governance with executive support of the Asset Management General Manager. It resulted in a data governance organisation based on the following roles and responsibilities:

- *Data Governance Council.* Membership of this council consists of executives from various divisions who have an interest in the management of asset data. They are responsible for endorsing policies, resolving cross divisional issues, engaging the IT council at the strategic level, strategically aligning business and IT initiatives, and reviewing budget submission for IT and non IT related projects.
- *Data Custodian.* Asset data is managed by the data custodian on behalf of Water Inc. He or she is responsible and accountable for the quality of asset data. The data custodian resolves issues raised in user group meetings. If issues become political and impacts stakeholders from other divisions, they are escalated to the DG council level. The data custodian is also responsible for endorsing data management plan, endorsing data cleansing plan, ensuring data is fit for purpose, converting strategic plans into tactical plans, change management, and stakeholder management.

- *Data Steward.* Data stewards have detail knowledge of the business process and data requirements. At the same time they also have good IT knowledge to be able to translate business requirements into technical requirements. They are led by the data custodian and are responsible for carrying out the tactical plans. They act on behalf of the data custodian in stakeholder management, change management, asset related information systems management and project management. They manage user group meetings, train and educate data users.
- *User Groups.* Data stakeholders from various divisions are invited to the user group meetings. These key data stakeholders consist of people who collect the data, process and report off the data. Technical IT staff is also invited to these meetings so that their technical expertise is available during the meeting. This is also a venue where urgent operational data issues can be tabled. The data users are responsible for reporting any data related issues, requesting functionality that would help them collect data more efficiently, and specifying reporting requirements.

The data governance structure in Figure 1 shows the business engagement with IT at the strategic, tactical and operational levels. This level of engagement ensures that IT and business are kept informed and IT initiatives align with the business data governance objectives.

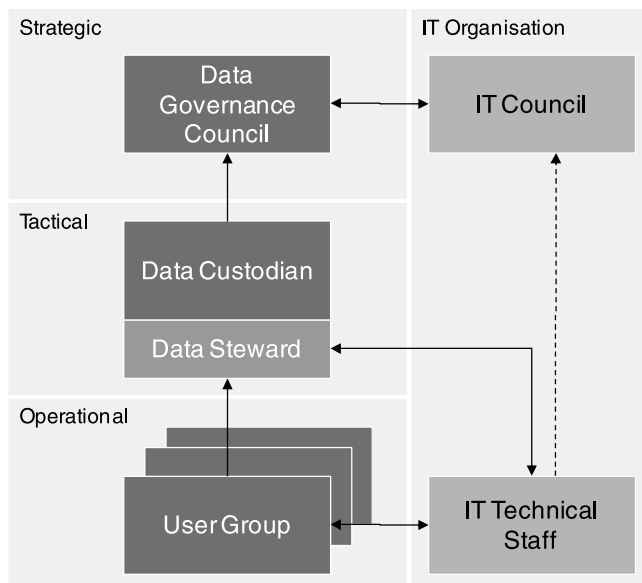


Figure 1: Data Governance Roles at Water Inc.

5 Discussion

The data governance structure at Water Inc. provides a structured framework for mitigating the risks of DQM. It is scalable to include other divisions as the data governance efforts within the DM team mature. The data custodianship and user groups structure can be adopted by other division with the data governance council acting as the 'organisational glue'. Water Inc. had shown that business and IT need to work together in order to manage corporate data effectively.

The roles and committee used in the data governance structure correspond on the whole to the set described in Section 2.2. The data governance council corresponds to the data quality board, the data custodian to the chief steward and the data stewards correspond to the business data stewards. The Asset Management General Manager is the executive sponsor of the data governance initiative. IT technical staff fills the role of the technical data steward. Because of the operational nature of the user groups they have no equivalent in a typical governance structure.

The DM team had provided insights into the difficulties in managing data as corporate asset without proper authority. The following summarises the findings of this research:

- *The justification for formal data governance.* This study had shown that managing the data quality of enterprise data is not effective without a formal data governance model. The reason for this is because of the lack of clear roles and responsibilities among data stakeholders. Data governance also assists business in engaging IT (vice versa) to manage corporate data collaboratively.
- *The process of setting up a formal data governance program.* The first step to setting up a formal data governance program is to determine a data governance structure. The structure provides escalation authority and a transparent decision-making process. Roles and responsibilities are defined so that members within the data governance structure are held accountable for their actions. Water Inc. had largely achieved this, however as this structure was recently introduced the success of its implementation cannot be determined in this study.
- *Ability to carry out actions as a result of a formal data governance structure.* Given the clear structure, the DM team is able to purchase a data profiling and metadata repository tool. The data profiling tool will allow the DM team to discover anomalies more efficiently. The metadata repository tool captures information about data so that it can be accessed by the whole organisation. Metrics measuring the quality of data had also been developed. The publication of these metrics will help management to determine the success of data improvement initiatives.
- *Simple data governance structure and framework.* The DM team did not want a cumbersome framework that could cause bottlenecks and delays in existing and future projects. With a proper structure and framework, the DM team is able to steer strategic projects to conform to and maintain good data governance. Nevertheless, Water Inc. moved from the governance via stewardship to the governance via governance approach.

6 Conclusion and Future Work

Companies need data quality management that combines business-driven and technical perspectives to respond to strategic and operational challenges demanding high-quality corporate data. Data governance specifies the framework for decision rights and accountabilities as part of corporate-wide DQM.

This paper investigates whether effective data quality management can be achieved without formal data governance. Some insights into data governance initiatives in a large utility organisation were obtained from this study. This paper underscores the importance of a data governance structure together with policies and procedures for managing data effectively. A data governance framework also enables collaboration from various levels of the organisation and it also provides the ability to align various data related programs with corporate objectives. This paper highlights that data governance provides a structured framework for mitigating the risks of data management.

This research has thrown up many questions in need of further investigation. Given that Water Inc. had only recently introduced a data governance structure, a longitudinal study would give a better indication of the benefits and success of the data governance structure implementation. This case study was conducted in isolation. Future research could investigate multiple case studies and compare the implementation of data governance programs between organisations of different sizes.

With this paper we encourage IT and business to work together to achieve high quality data. The case study had shown that IT has the technical know-how but business sponsorship is important to give data quality programs visibility and direction.

7 References

- [BGM87] Benbasat, I.; Goldstein, D. K.; Mead, M.: The Case Research Strategy in Studies of Information Systems. In: *MIS Quarterly* 11 (1987) 3, pp. 369-386.
- [BN07] Bitterer, A.; Newman, D.: *Organizing for Data Quality*. Gartner Research, Stamford, CT 2007.
- [CM06] Crié, D.; Micheaux, A.: From Customer Data to Value: What is lacking in the information chain? In: *Database Marketing & Customer Strategy Management* 13 (2006) 4, pp. 282-299.
- [CS01] Clemmons, S.; Simon, S.: Control and Coordination in Global ERP configuration. In: *Business Process Management Journal* 7 (2001) 3, pp. 205-215.
- [De06a] Dember, M.: 7 Stages for Effective Data Governance. In: *Architecture & Governance Magazine* 2 (2006) 4.
- [De06b] Dember, M.: Case Study: Gaining Efficiencies as a Result of Implementing a Data Governance Program. Presentation 2006.
- [DL06a] Dyché, J.; Levy, E.: *Customer Data Integration*. John Wiley & Sons 2006.
- [DL06b] Dyché, J.; Levy, E.: *Checkfree*. John Wiley & Sons, Hoboken, New Jersey 2006.

- [Ei89] Eisenhardt, K. M.: Building Theories form Case Study Research. In: *Academy Of Man-agement Review* 14 (1989) 4, pp. 532-550.
- [En99] English, L. P.: *Improving Data Warehouse and Business Information Quality*. 1 edition, John Wiley & Sons, Inc., New York, NY 1999.
- [Ep06] Eppler, M. J.: *Managing Information Quality*. 2nd Edition, Springer, Berlin, Heidelberg 2006.
- [Fr06] Friedman, T.: *Gartner Study on Data Quality Shows That IT Still Bears the Burden*. Gartner Group, Stamford 2006.
- [HLW99] Huang, K.-T.; Lee, Y. W.; Wang, R. Y.: *Quality Information and Knowledge*. Prentice Hall, New Jersey 1999.
- [KÖ06] Kagermann, H.; Österle, H.: *Geschäftsmodelle 2010 – Wie CEOs Unternehmen transformieren*. Frankfurter Allgemeine Buch, Frankfurt 2006.
- [Ma04] Madnick, S. et al.: *Exemplifying Business Opportunities for Improving Data Quality through Corporate Household Research*. MIT Sloan School of Management, Cambridge, MA 2004.
- [MS06] Marco, D.; Smith, A. M.: *Metadata Management & Enterprise Architecture: Understanding Data Governance and Stewardship*. In: *DM Review* (2006) Sep/Okt/Nov.
- [NL06a] Newman, D.; Logan, D.: *Achieving Agility: How Enterprise Information Management Overcomes Information Silos*. Gartner Research, Stamford, CT 2006.
- [NL06b] Newman, D.; Logan, D.: *Governance Is an Essential Building Block for Enterprise Information Management*. Gartner Research, Stamford, CT 2006.
- [PS05] Price, R.; Shanks, G.: *A Semiotic Information Quality Framework: Development and Comparative Analysis*. In: *Journal of Information Technology* 2005 (2005) 20, pp. 88-102.
- [RC05] Reid, A.; Catterall, M.: *Invisible data quality issues in a CRM implementation*. In: *Journal of Database Marketing & Customer Strategy Management* 12 (2005) 4, pp. 305-314.
- [Re96] Redman, T. C.: *Data Quality for the Information Age*. Artech House, Boston, London 1996.
- [Ru06] Russom, P.: *Taking Data Quality to the Enterprise through Data Governance*. The Data Warehousing Institute, Seattle 2006.
- [Sw05] Swanton, B.: *Master Data Management Organizations: A Balance of Control and Responsibility*. AMR Research, Boston 2005.
- [SZW03] Shankaranarayan, G.; Ziad, M.; Wang, R. Y.: *Managing Data Quality in Dynamic Decision Environments: An Information Product Approach*. In: *Journal of Database Management* 14 (2003) 4, pp. 14-32.
- [Te04] Tellkamp, C. et al.: *From Pallet to Shelf: Improving Data Quality in Retail Supply Chains Using RFID*. In: *Cutter IT Journal* 17 (2004) 9, pp. 19-24.
- [Th06] Thomas, G.: *Alpha Males and Data Disaster – The Case for Data Governance*. Brass Canon Press, USA 2006.

- [Ve00] Vermeer, B. H. P. J.: How Important is Data Quality for Evaluating the Impact of EDI on Global Supply Chains? In: Proceedings, Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (HICSS-33) 2000.
- [Wa98a] Wang, R. Y. et al.: Manage Your Information as a Product. In: Sloan Management Review 39 (1998) 4, pp. 95-105.
- [Wa98b] Wang, R. Y.: A Product Perspective on Total Data Quality Management. In: Communications of the ACM 41 (1998) 2, pp. 58-65.
- [We04] Weill, P.: Don't just lead, govern: How top-performing firms govern IT. In: MIS Quarterly Executive 3 (2004) 1, pp. 1-17.
- [WFA04] Watson, H. J.; Fuller, C.; Ariyachandra, T.: Data warehouse governance: best practices at Blue Cross and Blue Shield of North Carolina. In: Decision Support Systems 38 (2004) 3, pp. 435-450.
- [WO07] Wende, K.; Otto, B.: A Contingency Approach to Data Governance. In: Proceedings, 12th International Conference on Information Quality (ICIQ-07), Cambridge, USA 2007.
- [Yi02] Yin, R. K.: Case Study Research: Design and Methods. 3rd Ed. edition, Sage Publications, Thousand Oaks 2002.
- [ZG03] Zahay, D.; Griffin, A.: Information antecedents of personalisation and customisation in business-to-business service markets. In: Journal of Database Marketing 10 (2003) 3, pp. 255-271.

Integration und Integrationsmanagement

Leitung und Organisation des Tracks

N. Gronau, Universität Potsdam

Programmkomitee des Tracks

S. Aier, Universität St. Gallen
U. Bub, Telekom Laboratories
J. vom Brocke, Hochschule Liechtenstein
N. Gronau, Universität Potsdam
O. Höft, O2 Germany
T. Hübsch, Deutsche Telekom
P. Loos, Universität Saarbrücken
W. Pussak, Bayer Business Services
J. Rick, Deutsche Bahn AG
J. Schelp, Universität St. Gallen
M. Schönherr, Deutsche Telekom Laboratories
K. Weronek, Fraport AG
M. Weske, Hasso-Plattner-Institut

Szenariobasierte Entwicklung der Unternehmensarchitektur bei der Real-Time Center AG

Stephan Aier*, Nissim J. Buchs**, Jan Saat*, Guido Steiner**

*Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedberg-Strasse 8
CH-9000 St. Gallen
{stephan.aier | jan.saat}@unisg.ch

**RTC Real-Time Center AG
Schwarzenburgstrasse 160
CH-3097 Liebefeld
{nissim.buchs | guido.steiner}@rtc.ch

Abstract: In diesem Beitrag wird die szenariobasierte Entwicklung der Unternehmensarchitektur vorgestellt. Theoriegeleitet und auf bewährten Ansätzen aufbauend wird hierzu ein Vorgehen entwickelt, welches den Anwendungszweck der beteiligten Stakeholder in den Mittelpunkt der Metamodellentwicklung stellt. Im Anschluss wird der Einsatz dieses Vorgehens in einem Umsetzungsprojekt bei der Real-Time Center AG beschrieben. Abschließend wird das Vorgehen in einem kritischen Fazit beurteilt. Darüber hinaus werden nächste Schritte der Projektarbeit und Potenziale künftiger Forschung aufgezeigt.

1 Einleitung

Die Beweggründe für eine Organisation, sich mit dem Thema Unternehmensarchitekturen zu beschäftigen, sind vielfältig. Häufig entsteht im Laufe der Zeit eine heterogen gewachsene Informationssystemlandschaft, die eine nur schwer zu beherrschende Komplexität mit sich bringt. Die Folge sind permanente Aufwände, die für die Bewahrung beziehungsweise Wiederherstellung von Konsistenz in den Teilarchitekturen aufzubringen sind. Weitere Aufwände entstehen für den Abgleich fachlicher Anforderungen mit den Strukturen der Informationssysteme.

Der Schwerpunkt eines Architekturmanagements kann im Management der Abhängigkeiten der verschiedenen Artefakttypen der Unternehmensarchitektur gesehen werden. Dieses Abhängigkeitsmanagement trifft oft auf ein unterschiedlich stark ausgeprägtes Management der jeweiligen Artefakte eines Artefakttyps. Diese dezentralen Managementansätze haben oft eigene Begriffssysteme, Denkwelten und Metamodelle hervorgebracht.

In der Konsequenz sind standardisierte Unternehmensarchitekturmodelle oder Frameworks wie das Zachmann-Framework [Za87], FEA [CI99] oder TOGAF [To07] zwar gute Testmodelle, um die Konsistenz und Vollständigkeit des eigenen Verständnisses zu überprüfen, jedoch können sie nur selten unverändert auf eine Organisation angewendet werden. Signifikant wird dieses Problem, wenn zur Unterstützung des Architekturmanagements standardisierte Software-Werkzeuge eingesetzt werden sollen, die ein festes Metamodell zu Grunde legen. Hierbei treten vielfältige Risiken auf. So werden beispielsweise bestehende Begriffssysteme destabilisiert. Die Folge sind Akzeptanzprobleme für das Projekt und das einzuführende Werkzeug. Weiterhin besteht die Gefahr, dass vielfältige Modelltypen der Werkzeuge genutzt werden, ohne dass die Einsatzszenarien, die daran beteiligten Stakeholder und deren Informationsbedarfe verstanden worden sind. Außerdem besteht die Gefahr, dass hohe Aufwände in die initiale Erstellung der oft zu detaillierten Architekturmodelle fließen und deren Wartung nach Einführung nicht gewährleistet werden kann. Da veraltete Modelle keinen Nutzen erzeugen, scheitern einige Umsetzungsprojekte unmittelbar nach der Ersteinführung.

Um die genannten Risiken zu adressieren, sollte der Aufbau einer Unternehmensarchitektur von den verfolgten Einsatzszenarien und den dabei auftretenden, individuellen Informationsbedarfen der relevanten Stakeholder geleitet werden.

Der vorliegende Beitrag zeigt, wie ein solches szenariobasiertes Vorgehen zur Entwicklung einer Unternehmensarchitektur bei der Real-Time Center AG (RTC) angewendet wurde und richtet sich somit sowohl an Wissenschaftler als auch an Praktiker. Im folgenden Abschnitt 2 werden einige theoretische Grundlagen zu Unternehmensarchitekturen vorgestellt. In Abschnitt 3 wird unter Verwendung verwandter Ansätze ein szenariobasiertes Vorgehen entwickelt, dessen Anwendung anhand einer Projektbeschreibung in Abschnitt 4 dargestellt ist. Abschnitt 5 zeigt Erfahrungen und Anwendungsgrenzen des methodischen Vorgehens sowie Besonderheiten des Projektes auf. Im Zuge dieser kritischen Reflexion werden weitere Projektschritte sowie Potenziale für Forschungsaktivitäten vorgeschlagen.

2 Theoretische Grundlagen

Gemäß ANSI/IEEE Standard 1471-2000 ist Architektur definiert als “the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution” [IE00]. Die Unternehmensarchitektur stellt in Folge die fundamentale Strukturierung einer Organisation (Unternehmen, Behörde etc.) dar.

In der Regel wird die Unternehmensarchitektur nicht durch eine einzige Architektur, sondern durch eine Reihe von Teilarchitekturen beschrieben, die bspw. in Ebenen organisiert werden können [To07; WF07; Za87]. Die Unternehmensarchitektur vernetzt also bestehende Modelle, wie etwa Datenmodelle, Geschäftsprozessmodelle, Komponentenmodelle, Modulstrukturdiagramme, Netzwerkpläne, Inventarlisten, Hardwarelisten sowie Organigramme, Arbeitsplatzbeschreibungen oder Funktionsbäume zu einem konsistenten, umfassenden Abbild der Unternehmung. Hiervon ausgehend besitzt jedes

Unternehmen eine Unternehmensarchitektur, nur wird sie nicht in jedem Unternehmen genutzt [Ni05]. Kernartefakt der Unternehmensarchitektur ist ein Metamodell, welches die Quintessenz des Gestaltungsgegenstandes „Unternehmen“ abbildet [OW07]. In einem Metamodell werden die Artefakttypen der Teilarchitekturmodelle abgebildet und zueinander in Beziehung gesetzt [KL05; WF07]. Die Unternehmensarchitektur als Abbild der Organisation fungiert somit als Kommunikationsbasis und Dokumentation des Ist-Zustandes der Unternehmensstrukturen und Prozesse [Pa06], die Unterstützung für die Erstellung des Soll-Zustandes der Unternehmensstrukturen und Prozesse sowie die Unterstützung von Transformationsprojekten vom Ist- zum Soll-Zustand [WB07].

Des Weiteren sollen durch die Unternehmensarchitektur fachliche Analysen ermöglicht und somit interdisziplinäre Fragen beantwortet werden. Beispiele für derartige Analysen und Auswertungen sind [Ni05]:

- Abhängigkeits-/Auswirkungsanalyse (beispielsweise Analyse der Beziehungen zwischen den Artefakten, Analyse der Auswirkung von Änderungen eines Artefaktes auf andere Artefakte)
- Abdeckungsanalyse (beispielsweise Analyse der Abdeckung fachlicher Bereiche durch Anwendungssysteme)
- Schnittstellenanalyse (beispielsweise Analyse der Schnittstellen innerhalb eines Artefakttyps, etwa Art, Anzahl, Komplexität, Häufigkeit, Stabilität)
- Heterogenitätsanalyse (Analyse der Heterogenität, beispielsweise Analyse der Entwicklungslinien in definierten Einsatzfeldern)
- Komplexitätsanalyse (Analyse der Komplexität der Unternehmensarchitektur, beispielsweise Analyse der Komponenten und ihrer Beziehungen)
- Konformitätsanalyse (beispielsweise Analyse der Einhaltung von Standards oder Compliance Rules und Ermittlung des Abweichungsgrads)
- Kostenanalyse (beispielsweise Analyse der Erstellungs-, Betriebs- und Wartungskosten)
- Nutzenanalyse (Analyse des Nutzens, beispielsweise Verfolgung der Erreichung von Unternehmenszielen, Verfolgung von Kennzahlen)

Als besonders kritisch wird häufig die Schnittstelle zwischen der IT und den Fachbereichen empfunden. Kernfragen, die in diesem Kontext durch die Unternehmensarchitektur beantwortet werden sollen, lauten beispielsweise: Welche Geschäftsprozesse werden durch (welche) IT-Systeme unterstützt? Wie und womit wird diese Unterstützung realisiert? Welche Kosten und welcher Nutzen entstehen hierdurch? [Ni05]

3 Einführung eines szenariobasierten Vorgehens zur Entwicklung der Unternehmensarchitektur

Durch ihren holistischen Charakter gilt die Unternehmensarchitektur auch als Management-Informationssystem des CIO [CI05; JL06; LJ06; SL06]. Es gibt eine Vielzahl denkbarer Einsatzszenarien mit Stakeholdern aus der gesamten Organisation. Beispielhafte Einsatzszenarien sind in Tabelle 1 aufgelistet [vgl. auch WB07].

Einsatzszenario	Quelle
Business Continuity Planning	[Ke07]
Produktplanung	[AO05; TC06]
Projektportfolioplanung/Projektselektion	[Ke07; WB07]
Business/IT-Alignment	[La05; Lu05; Pa06; PS05; SL06]
Security Management	[WB07]
Technology Risk Management	[WB07]
Prozessoptimierung	[JL06; Pa06]
Qualitätsmanagement	[WB07]
Compliance Management	[JL06; Ke07]
Konfigurationsmanagement	[AN98; Kr05]
Post-Merger-Integration	[WB07]
Standardsoftware-Einführung	[WB07]
Unterstützung von Sourcing-Entscheidungen	[WB07]
IT-Service-Management	[BW07; MB05]
IT-Betriebskostenoptimierung	[CI05; Ke07]
IT-Konsolidierung	[CI05]

Tabelle 1: Einsatzszenarien der Unternehmensarchitektur

Ein Einsatzszenario der Unternehmensarchitektur lässt sich über die Anforderungen der beteiligten Stakeholder charakterisieren. Hierbei wird davon ausgegangen, dass eine Person, eine Organisationseinheit oder eine Rolle (ein Stakeholder) eines Einsatzszenarios bestimmte „Concerns“ in Bezug auf die zu bewältigenden Aufgaben hat. In Anlehnung an [IE00] werden Concerns vorliegend als diejenigen Interessen der Stakeholder verstanden, die die Entwicklung, den Betrieb oder einen anderen für einen oder mehrere Stakeholder relevanten oder kritischen Aspekt der Unternehmensarchitektur betreffen. Determiniert durch die Concerns hat ein Stakeholder Informationsbedarfe. Das Metamodell der Unternehmensarchitektur besteht aus Gestaltungsobjekten und deren Verbindungen. Wenn das Modell in der Organisation instanziiert ist und zu den Gestaltungsobjekten Daten hinterlegt wurden, lassen sich idealerweise durch Abfragen und Analysen der Gestaltungsobjekte und deren Verbindungen die Informationsbedarfe der Stakeholder befriedigen. Hierzu müssen die Informationen aus dem Modell zielgruppenadäquat dargestellt werden. Abbildung 1 verdeutlicht den Sachverhalt.

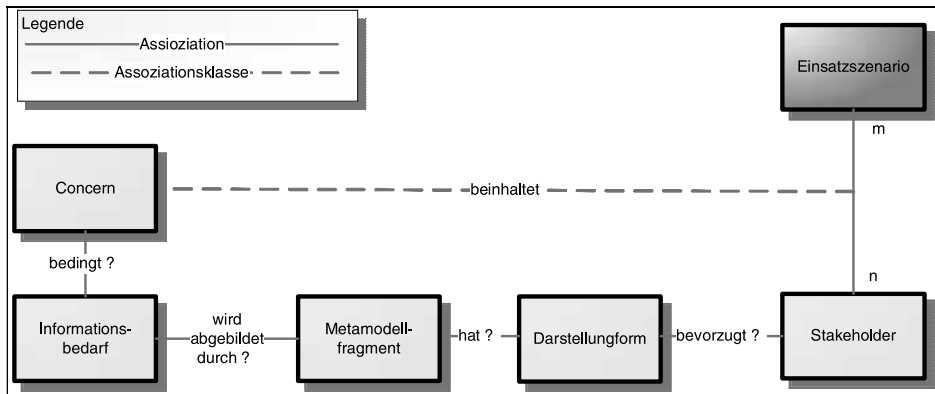


Abbildung 1: Metamodell des Einsatzszenarios

Angelehnt an [KW07] und ausgehend von der These, dass ein Metamodell die Summe aller Informationsbedarfe der Concerns aller Stakeholder der Unternehmensarchitektur befriedigt, kann ein Metamodell nach folgendem Vorgehen entwickelt werden.

- Schritt 1) Definition der Einsatzszenarien
- Schritt 2) Identifikation der Stakeholder
- Schritt 3) Stakeholderinterviews
- Schritt 4) Metamodellierung

Das Ergebnis dieser vier Schritte ist ein Metamodell der Unternehmung, dessen Instanziierung die Informationsbedarfe der Stakeholder abdeckt, die an den Einsatzszenarien beteiligt sind. Zur Umsetzung der Unternehmensarchitektur sind weitere Handlungen notwendig, die sich einem pragmatischen Vorgehen folgend in vier weitere Schritte untergliedern lassen:

- Schritt 5) Toolauswahl
- Schritt 6) Iterative Umsetzung
- Schritt 7) Pflege und Betrieb
- Schritt 8) Kommunikation

Während die Schritte 1) bis 4) die Entwicklung des Metamodells gliedern, ist die konkrete Umsetzung und der kontinuierliche Betrieb der Unternehmensarchitektur Gegenstand der Schritte 5) bis 8). Der Fokus dieses Beitrages liegt auf den Schritten 1) bis 4).

4 Anwendung des Vorgehens bei der RTC

In den folgenden Abschnitten wird der Aufbau der Unternehmensarchitektur bei der RTC beschrieben. Zunächst werden in 4.1 einige Informationen zum Unternehmen und zur Ausgangslage des Projektes gegeben. Abschnitt 4.2 stellt die einzelnen Projektschritte aus Abschnitt 3 für den Kontext des Unternehmensarchitekturprojektes bei der RTC detailliert dar.

4.1 Vorstellung der RTC und Motivation zur Entwicklung der Unternehmensarchitektur

Die Real-Time Center AG (RTC) mit Sitz in Bern Liebefeld wurde im Jahre 1973 mit dem Auftrag gegründet, eine Softwareplattform für Banken zu entwickeln und diese im Outsourcing für interessierte Kundenbanken zu betreiben. Es entstand die Bankenplattform IBIS® für Universal-, Retail- und Vermögensverwaltungsbanken. Diese wurde erweitert und das Angebot an Outsourcing-Dienstleistungen ausgebaut. Seit mehr als 30 Jahren bietet RTC Softwareprodukte und IT-Outsourcing-Lösungen an. RTC verfügt über Erfahrung im Betrieb von Serverfarmen sowie in der Konzeption und dem Management von komplexen Netzwerken, Hochsicherheitslösungen und dezentralen Microsoft-Umgebungen. Archivierungs- und Backup-Dienstleistungen, das Druck- und Versandzentrum sowie Ausbildungs- und Helpdesk-Dienstleistungen ergänzen das Angebot. RTC beschäftigt 550 Software- und Outsourcing-Spezialisten, verfügt über zwei eigene Data Center und ist schweizweit in allen Sprachregionen tätig. Zu den Kunden zählen über 50 Banken mit mehr als 600 Vertriebsstellen, welche über 4 Mio. Bankkunden betreuen [RT08].

Die RTC verfügt über zahlreiche detaillierte, jedoch isolierte Bestandsführungen und Teilarchitekturen. Es fehlt ein durchgängiges Gesamtbild der Unternehmung. Ziel ist es aus diesem Grund, eine vollständige Unternehmensarchitektur in Form einer Gesamtdokumentation vom Geschäftsmodell bis hin zur Infrastruktur zu entwickeln. Hierbei wird das Unternehmen nicht von Grund auf neu abgebildet, sondern baut auf bestehenden Modellen (Bestandsführungen) und deren Integration auf. Ziel ist, die entstehende Gesamtarchitektur als Steuerungsinstrument und Analysegegenstand nutzen zu können um mehr Transparenz und erhöhte Flexibilität im Unternehmen zu erreichen. Hierzu können die in Abschnitt 2 vorgestellten Analysen dienen.

4.2 Szenariobasiertes Vorgehen

Im Folgenden wird die Anwendung des oben dargestellten Vorgehens bei der RTC diskutiert. Bei der praktischen Anwendung zeigt sich, dass die einzelnen Schritte nicht strikt nacheinander, sondern größtenteils parallel ablaufen. Weiterhin wird im beschriebenen Anwendungsfall ein bestehendes Metamodell weiterentwickelt und nicht, wie bei [KW07] vorgestellt, ein Modell von Grund auf neu entwickelt.

Zum Beschreibungszeitpunkt wurden die Schritte 1) bis 5) bereits durchlaufen beziehungsweise abgeschlossen. Die Schritte 6) bis 8) folgen in weiteren Teilprojekten respektive im kontinuierlichen Betrieb, weshalb hierzu in diesem Beitrag lediglich Ideen, jedoch keine Erfahrungen beschrieben werden.

Schritt 1) Definition der Einsatzszenarien

Da die Unternehmensarchitektur nicht simultan für alle Einsatzszenarien parallel entwickelt werden kann, muss eine Auswahl für initiale Projekte erfolgen. Ausgehend von den in Abschnitt 4.1 erläuterten Treibern und Zielen des Projektes wurden die hierfür wichtigsten Einsatzszenarien der Unternehmensarchitektur identifiziert. Hierzu wurden die Elemente aus Tabelle 1 vor dem Hintergrund der Zielsetzung priorisiert. Als die zunächst umzusetzenden Szenarien wurden Business/IT-Alignment, Konfigurationsmanagement und Compliance Management benannt.

Business/IT-Alignment wird häufig als wichtigstes Thema von IT-Führungskräften dargestellt [Lu05]. „Business-IT alignment refers to applying Information Technology (IT) in an appropriate and timely way, in harmony with business strategies, goals and needs.“ [Lu00] Hierzu ist nicht nur die Harmonisierung der Zielsysteme von Fachseite und IT erforderlich sondern auch ein holistisches Verständnis der gesamten Unternehmung.

Dem Compliance Management, also dem Befolgen von Gesetzen und Regeln [Ke07], kommt insbesondere bei Finanzdienstleistern eine Schlüsselrolle zu, da die Sicherheitsanforderungen stetig steigen. Gesetze, wie etwa das Schweizer Bundesgesetz über die Banken und Sparkassen (BankG) [Ba06], der Sarbanes-Oxley Act (SOX) [SO02] und Basel II [Bs06], haben zum Ziel, die Sicherheit des Finanzsystems zu erhöhen. Hiermit eng verbunden ist das Bestreben nach klaren Strukturen in der Unternehmensarchitektur, beispielsweise in Bezug auf Reporting, Daten- und Systemverantwortung und dedizierte Sicherheitsaspekte, wie etwa Authentifizierungs- und Autorisierungskonzepte [Wo06].

Unter Konfigurationsmanagement wird „...a management process for establishing and maintaining consistency of a product's performance, its functional and physical attributes, with its requirements, design and operational information, throughout its life“ [AN98] verstanden. Allgemeiner formuliert besteht Konfigurationsmanagement aus der Administration und der Konfiguration sämtlicher IT-Komponenten im Unternehmen [Kr05]. Als Outsourcing-Dienstleister kommt der strukturierten Verwaltung und Steuerung von Hardware und Software in der RTC besondere Aufmerksamkeit zu, da eine hohe Transparenz der Konfigurationseinheiten jederzeit Voraussetzung für die Anbindung oder Ausgliederung von Kunden, Integration oder Entkopplung von Partnern sowie die Leistungsverrechnung ist.

Aufgrund der Brisanz des Themas Konfigurationsmanagement wurde zu Projektbeginn definiert, dass das entstehende Metamodell der Unternehmensarchitektur auch als Datenmodell für den Aufbau einer Konfigurationsmanagementdatenbank (CMDB) dienen sollte.

Schritt 2) Identifikation der Stakeholder

In einem nächsten Schritt wurden die Stakeholder der Unternehmensarchitektur der drei identifizierten Einsatzszenarien benannt. Hierbei wurden der dokumentierten Aufbauorganisation der RTC folgend und unter Berücksichtigung von Verantwortlichkeiten in derzeit laufenden Projekten diejenigen Rollen identifiziert, deren Anforderungen in Bezug auf die Einsatzszenarien in dieser ersten Iteration der Umsetzung berücksichtigt werden sollten.

	Business/IT-Alignment	Konfigurationsmanagement	Compliance Management
Stakeholder	Architektur	Change Management	Internes Kontrollsystem
	Produktmanagement	Release Management	Security
	Account Management	IT-Service Management	
	Applikationsverantwortliche		

Tabelle 2: Einsatzszenarien und Stakeholder

Wie in Tabelle 2 dargestellt, wurden insgesamt neun Rollen als Stakeholder der oben genannten Einsatzszenarien ermittelt. Mit den identifizierten Stakeholdern wurden in einem Folgeschritt Interviews durchgeführt.

Schritt 3) Stakeholderinterviews

Die Interviewtermine wurden jeweils durch eine Präsentation mit Hintergrundinformationen zu Unternehmensarchitekturen und dem damit verbundenen Potenzial eröffnet. Weiterhin wurden Informationen zum Projekt und dessen Teilschritten gegeben. Die Stakeholderbefragung erfolgte auf Basis eines Interviewleitfadens. In Anlehnung an das von [KW07] entwickelte Viewpoint-Requirement-Template wurden die Stakeholder in strukturierten Interviews nach Concerns, Informationsbedarfen, Metamodellfragment und derzeitiger vorhandener sowie künftig gewünschter Darstellungsform der von ihnen benötigten Informationen befragt (vgl. Tabelle 3).

Thema	Fragen
Concerns	Welche Aufgaben nehmen Sie innerhalb der RTC wahr? Wie könnte eine durchgehende Unternehmensarchitektur Sie in Ihren Aufgaben unterstützen?
Informationsbedarf	Was sind derzeit unzureichend beantwortete Fragen zu deren Klärung die Unternehmensarchitektur beitragen könnte?
Benefits	Was sind die (qualitativen) Benefits für Sie/Ihre Organisationseinheit/RTC, wenn Sie über die genannten Informationen verfügen würden?
Darstellungsform IST	Falls die o.g. Informationen derzeit für Sie schon zugänglich sind, wie werden diese dargestellt? Bzw. auf welchem Weg können Sie diese Informationen beschaffen?
Darstellungsform SOLL	Wie könnte die Bereitstellung, Aufarbeitung und Darstellung der benötigten Informationen ausgestaltet bzw. verbessert werden?

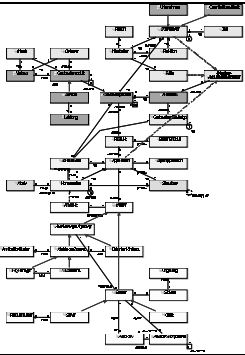
Metamodellfragment	Sind im derzeitigen Stand des RTC-Metamodells alle Gestaltungsobjekte (ggf. mit den jeweils benötigten Attributen) und Verbindungen (ggf. mit den jeweils benötigten Kardinalitäten) abgebildet, um die o.g. Informationsbedarfe zu decken?	
--------------------	---	--

Tabelle 3: Strukturierungsthemen der Stakeholderinterviews zur Anforderungserhebung

Um möglichst umfassende Informationen aus den Interviews generieren zu können, wurden die Fragen sehr offen gestellt. Als Konsequenz wurden die Antworten eher unscharf formuliert. In einem nächsten Teilschritt erfolgte die Transformation der Stakeholderwünsche in prüfbare Anforderungen. Tabelle 4 stellt auszugsweise einige Beispiele der erhobenen Anforderungen dar.

Req_ID	Kategorie	Beschreibung	Quelle
RQ0001	CMDB/EA	Die Unternehmensarchitektur beinhaltet alle Referenzbankprozesse (BEKB) für IBIS.	ITBA-AccMgmnt
RQ0003	CMDB/EA	Prozesse sind so feingranular definiert, dass jeder Aktion eindeutig eine Applikation oder eine Rolle zugewiesen werden kann.	ARLA
RQ0002	EA	Die Prozesse sind gemäß ROEP Geschäftsprozess Modellierungsdisziplin modelliert.	ITBA-AccMgmnt
RQ0004	CMDB	Jeder angebotene RTC-Service hat einen definierten Preis.	ITBA-AccMgmnt
RQ0005	CMDB/EA	Jede Ressource, die in einen RTC-Service einfließt, verursacht definierte Kosten.	ITBA-ProdMgmnt
RQ0006	CMDB/EA	Die Beziehung zwischen RTC-Service und SLA wird abgebildet.	ITBA-AccMgmnt
RQ0007	CMDB/EA	Die Beziehung zwischen Ressource und RTC-Service wird abgebildet.	ITBA-ProdMgmnt

Tabelle 4: Anforderungen an die Unternehmensarchitektur (Auszug)

Schritt 4) Metamodellierung

Die Metamodellierung der Unternehmensarchitektur wird in der RTC als kontinuierlicher Prozess verstanden. Das Metamodell ist in UML 2.0 dargestellt und umfasst die Teilarchitekturen Geschäftsarchitektur, Applikationsarchitektur, Softwarearchitektur und Betriebsarchitektur. Die Geschäftsarchitektur umfasst Gestaltungsobjekte wie beispielsweise Geschäftsprozess, Kunde, Lieferant, Mitarbeiter, Rolle, Organisation, Leistung und Vertrag (Service Level Agreement). Die Applikationsarchitektur beinhaltet beispielsweise IBIS-Applikation, IBIS-Produkt, Komponente und Schnittstelle, während die Softwarearchitektur Gestaltungsobjekte wie etwa Klasse, Paket, Subsystem und Modul enthält. In der vierten Teilarchitektur, der Betriebsarchitektur, werden unter anderem Server, Schiene, Knoten, Netzwerk, Middleware, Datenbank und Application Cluster abgebildet.

Das Metamodell kann aus unterschiedlichen Beweggründen heraus weiterentwickelt werden. Zum einen werden Begriffsharmonisierungen umgesetzt, wenn es innerhalb der RTC zu Unklarheiten oder Doppelbezeichnungen kommt oder unklar ist, in welchem Zusammenhang ein Begriff zu anderen Begriffen steht. Eine dedizierte Arbeitsgruppe zur Begriffsharmonisierung ist hierbei für die Definition, Umsetzung im Modell und Veröffentlichung im Glossar per Intranet verantwortlich. Zum anderen werden im Modell neu erhobene Anforderungen an die Unternehmensarchitektur und damit auch an das Konfigurationsmanagement umgesetzt.

In einem ersten Schritt der Weiterentwicklung des RTC-Metamodells wurde dieses an diverse Standards angepasst. In den Bereichen Geschäfts- und Applikationsarchitektur erfolgte zunächst ein strukturierter Vergleich zu dem von [OW07] vorgestellten Referenzmodell, dem „Core Business Metamodel“ der Universität St. Gallen. Es wurden die Gestaltungsobjekte und deren Beziehungen mittels eines Mappings verglichen und analysiert. Als Ergebnis konnten Teile des Referenzmodells in das RTC Metamodell übernommen werden. Weiterhin wurde das Metamodell im Bereich Geschäftsarchitektur dem Metamodell von Ericsson/Penker [EP00] angepasst. Im Bereich Softwarearchitektur wurde der UML 2.0 Standard und das J2EE Profil [OM05] übernommen und im Bereich Betriebsarchitektur wurde das Metamodell dem CIM Standard [CI08] angepasst.

Weiterhin wurde in der Modellentwicklung der Einsatz von stakeholderspezifischen Sichten diskutiert. „A viewpoint covers a number of concerns and defines the information associated with the concern in the meta model“ [Ba04]. Die Gestaltung der Sichten erfolgt darum in Anlehnung an Stakeholderperspektiven und deren Informationsbedarfen. Als IT-Outsourcing-Unternehmen für Banken positioniert sich die RTC als Dienstleistungsunternehmen für Dienstleistungsunternehmen. Diese Beziehung mit ihren Schnittstellen zwischen Lieferanten und der RTC, der RTC und Kunde (Bank), sowie zwischen Bank und Bankkunde erfordert besondere Berücksichtigung in der Gestaltung der Sichten auf das Metamodell. Wenn beispielsweise im Metamodell das Gestaltungsobjekt Geschäftsprozess definiert ist, kann damit je nach Sicht unterschiedliches gemeint sein, etwa der Bank-Geschäftsprozess, der RTC-Geschäftsprozess oder der Lieferanten-Geschäftsprozess.

Um ein konsistentes und normalisiertes Metamodell realisieren zu können, wurden die Gestaltungsobjekte Bank-Geschäftsprozess, RTC-Geschäftsprozess und Lieferanten-Geschäftsprozess als UML-Spezialisierungen des Gestaltungsobjektes Geschäftsprozess modelliert. In der Konsequenz entsteht ein zwar normalisiertes, jedoch relativ unübersichtliches konsolidiertes Metamodell, welches jedoch im Vergleich zur Modellierung und Pflege dreier fast identischer Modelle deutlich handhabbarer scheint. Die Einführung von Sichten soll in diesem Zusammenhang bei der Bewältigung der entstehenden Komplexität helfen. Das Metamodell wird somit nie als Ganzes präsentiert, sondern immer nur in denjenigen Ausschnitten, die für den Betrachter relevant sind. Ein Ausschnitt stellt immer nur eine bestimmte Sicht (Viewpoint) auf das Metamodell dar, in welcher die nicht benötigten Gestaltungsobjekte und Beziehungen ausgeblendet werden (siehe Viewpoint-Konzept in [IE00]).

Neben der Erstellung von Sichten durch Selektion von Gestaltungsobjekten sind stakeholder-spezifische Erweiterungen (so genannte Extensions) möglich [OW07]. Um die Konsistenz der Syntax dieser Erweiterungen sicherzustellen, existieren eine Reihe definierter Erweiterungsmechanismen. So können Erweiterungen beispielsweise durch Hinzufügen, Spezialisieren oder Beschreiben der Detailstruktur eines Gestaltungsobjekts (durch Aggregationsbeziehung zu neuen Detail-Gestaltungsobjekten) oder durch Beschreibung der Detailstruktur einer Assoziationsbeziehung (durch Hinzufügen von Gestaltungsobjekten oder Assoziationsklassen), durch Beschreiben der Detailstruktur einer Aggregationsbeziehung und durch Hinzufügen eines Aggregationsobjekts erfolgen.

Schritt 5) Toolauswahl

Gemäß [CJ07] empfiehlt Gartner die in Schritt 3) gesammelten Anforderungen durch zwei unterschiedliche Werkzeuge abzudecken. Anforderungen zu Konfigurationsmanagement und damit zusammenhängenden Analyseformen, die beispielsweise für das Business Continuity Planning, das Incident Management oder Root Cause-Analysen werden demnach am besten durch ein entsprechendes CMDB-Tool abgedeckt. Anforderungen betreffend Buiness/IT-Alignement und Compliance wird hingegen am besten durch ein darauf spezialisiertes Unternehmensarchitektur-Tool begegnet.

Eine genaue Auseinandersetzung mit den erhobenen Anforderungen ergab, dass 50 % der Anforderungen durch die Einführung einer CMDB abgedeckt werden können. Weitere 40 % der Anforderungen können durch die Einführung eines Unternehmensarchitektur-Tools umgesetzt werden und 10 % der Anforderungen betreffen ITIL-Prozesse, die indirekt auf der CMDB aufbauen oder nicht abgedeckt werden können.

Aus diesen Gründen wurden zwei Vorhaben gestartet. Zum einen werden im Operations-Bereich Werkzeuge für das Konfigurationsmanagement evaluiert. Fokus dieses Werkzeuges sollen Darstellung des IST-Zustandes und Inventarisierungsfunktionen sein. Ein weiteres Projekt befasst sich mit der Evaluierung und Einführung eines dedizierten Werkzeugs für die Unternehmensarchitektur. Mithilfe dieses Werkzeuges sollen insbesondere Planungs- und Steuerungsfunktionen umgesetzt werden.

Schritt 6) Iterative Umsetzung

Nach Auswahl und Einführung wird im Rahmen eines ersten Umsetzungsprojektes („Proof of Concept“) der Mehrwert der Lösung demonstriert. Für den weiteren Projektverlauf wird ein iteratives Vorgehen zur Umsetzung der Anforderungen geplant. Hierzu werden wiederum Anforderungen detaillierter erhoben. Weiterhin werden die zur Umsetzung notwendigen Datenquellen identifiziert und integriert. Auf Basis der Anforderungen wird nun die konkrete Ausgestaltung der Reports erfolgen.

Schritt 7) Pflege und Betrieb

Um eine hohe Datenqualität der Modelle in den Teilarchitekturen zu gewährleisten, ist eine starke Governance der Unternehmensarchitektur unerlässlich. [FA07] schlagen hierzu einen generischen Pflegeprozess vor. Um ein ganzheitliches Konzept zur Governance und Pflege der RTC Unternehmensarchitektur zu adaptieren, werden in noch folgenden Projektschritten Analysen der Quellsysteme sowie der Verantwortlichkeitsstrukturen durchgeführt werden.

Schritt 8) Kommunikation

Kommunikationsaufgaben der Unternehmensarchitektur bestehen beispielsweise in der Ausbildung der Stakeholder, in der Formulierung und Abgrenzung der bereitgestellten Leistungen sowie in der Vermittlung von Methodiken und Richtlinien. Da die Unternehmensarchitektur ein klassisches Schnittstellenthema besetzt, kommt der Kommunikation eine Schlüsselrolle zu. Im Rahmen eines Kommunikationskonzeptes wird in geplanten Teilprojekten erarbeitet, mit welchen Kommunikationstechniken welchen Stakeholdergruppen gegenübergetreten wird, um zugrunde liegende Kommunikationsziele zu erreichen. Impulse für das Vorgehen liefern unter anderem [La05].

5 Fazit und Ausblick

Das szenariobasierte Vorgehen zur Entwicklung der Unternehmensarchitektur bei der RTC und die iterative Umsetzung der erhobenen Anforderungen wurde von den Projektbeteiligten als eine geeignete Methode wahrgenommen, um schnell und effizient zu den benötigten Anforderungen an das Metamodell, die benötigten Sichten und entsprechende Toolunterstützung zu kommen.

Für den praktischen Einsatz musste das vorgeschlagene Vorgehen geringfügig modifiziert werden. So erfolgten die Schritte nicht nacheinander sondern teilweise parallel. Insbesondere die Metamodellierung stellte sich statt eines isolierten Projektschrittes als eine das Projekt kontinuierlich begleitende Aufgabe heraus. Dies liegt zum Teil daran, dass zu Beginn des Projektes bereits ein Metamodell vorlag, welches während der Schritte 1)-4) stetig weiterentwickelt werden konnte, während die genutzte Methode die Annahme zugrunde legt, dass das Metamodell von Grund auf neu entwickelt wird.

Eine weitere Besonderheit, die nicht unmittelbar in der zugrunde liegenden Methode berücksichtigt war, lag in der Beteiligung des Betriebs. So wird neben klassischen Planungs- und Steuerungsaufgaben der Unternehmensarchitektur ein weiterer Fokus, nämlich die Inventarisierung und Verwaltung von Konfigurationseinheiten, relevant. Das Konfigurationsmanagement stellt gemäß ITIL den aktuellen Stand der IT-Infrastruktur vom Service Level Agreement bis zur physischen Hardwarekomponente dar und erlaubt es, transparent und rekonstruierbar zu erklären, wie es zu diesem Stand gekommen ist. Ziel der Unternehmensarchitektur ist es hingegen, aufzuzeigen, wie sich der aktuelle Stand in Zukunft gemäß der Strategie nach verschiedenen Szenarien verändern soll und was die Auswirkungen dieser Veränderungen sind. Es bestehen folglich potenzielle

Abhängigkeiten zwischen den Betrachtungs- und Analysegegenständen. Welche Informationen der CMDB auf welchem Aggregationslevel letztlich architekturelevant sind, ist im Einzelfall durch die Stakeholder zu entscheiden. Die Erfahrung zeigt, dass sich durch die Zusammenarbeit zwischen Unternehmensarchitekturprojekt und CMDB-Aufbau (z. B. gemeinsame Interviewtermine bei Stakeholdern) einige Synergien erzielen lassen. Die vorgeschlagene Methode und auch der Gestaltungsgegenstand beinhalten jedoch einige Spezifika, die nicht gemeinsam erarbeitet werden können. Die Durchführung der Stakeholderinterviews brachte die Erkenntnis, dass die Qualität der erhaltenen Informationen für eine Erhebung prüfbarer Anforderungen nicht ausreichend war. Somit wurde es erforderlich, die eher vage formulierten Bedürfnisse in einem Zwischenschritt in Anforderungen umzuschreiben.

Ergebnis der Schritte 1) bis 4) ist ein konsolidiertes Metamodell. Die Darstellung dieses Modells ist zur Kommunikation mit Stakeholdern, die überwiegend einen fachlichen Hintergrund haben, nur bedingt geeignet. Zur Förderung der Akzeptanz wurde als erster Showcase ein Prototyp entwickelt, der die Kernzusammenhänge darstellt. Der Prototyp baut auf existierenden Stammdatenquellen ein Eclipse Modeling Framework auf und übersetzt dieses mittels Query-View-Transformation in ein UML 2.0-Modell. In einem letzten Schritt werden dann UML-Diagramme automatisch generiert, um so eine Vorschau auf die von den Auftraggebern gewünschten Informationen und Sichten zu liefern.

Eine weitere zentrale Initiative zur Förderung der Akzeptanz ist ein konsistentes Begriffsverständnis. Der Aufbau eines zentralen Glossars, in dem die modellrelevanten Artefakte definiert sind hat sich im beschriebenen Projekt bewährt. Die Projekterfahrung zeigt, dass dieser Schritt im Zuge der Metamodellierung erfolgen kann. Als kritischer Erfolgsfaktor für die nächsten Schritte wird insbesondere die konsequente Pflege der Modelldaten gesehen, da die Akzeptanz der Unternehmensarchitektur maßgeblich von der Datenqualität abhängt. Weiterer Forschungsbedarf liegt hier in der Identifikation typischer Szenarien und möglicher Pflege- und Governance-Strategien.

Weitere Forschungsaktivitäten ergeben sich aus dieser Arbeit insbesondere in Hinblick auf die Weiterentwicklung der Methode. Vorliegend wird davon ausgegangen, dass das Metamodell von Grund auf entwickelt wird. In vielen Fällen wird jedoch ein Referenzmodell Grundlage für die Unternehmensarchitektur sein. In diesem Zusammenhang sollten Adaptionmechanismen sowie die Erkenntnisse aus dem situativen Methodenengineering, die die Anwendung und Umsetzung erleichtern können, betrachtet werden. Weiterhin ist die Anschlussfähigkeit zu verwandten Ansätzen zu prüfen und sicherzustellen, so dass die Anwendung als Best-of-breed methodisch fundiert erfolgen kann.

6 Literaturverzeichnis

- [AN98] ANSI/EIA: 649-1998 National Consensus Standard for Configuration Management. (1998).
- [AO05] Aziz, S. et al.: Enterprise Architecture: A Governance Framework – Part I: Embedding Architecture into the Organization. Infosys, 2005.

- [Ba04] Bayer, J.: View-based Software Documentaion. Master Thesis, Universität Kaiserslautern, Kaiserslautern, 2004.
- [Ba06] BankG: Bundesgesetz vom 8. November 1934 über die Banken und Sparkassen. http://www.admin.ch/ch/d/sr/c952_0.html, Abruf am 11.03.2008.
- [Bs06] Basel-II: Die neue Baseler Eigenkapitalvereinbarung. http://www.bundesbank.de/bankenaufsicht/bankenaufsicht_basel.php, Abruf am 11.03.2008.
- [BW07] Braun, C.; Winter, R.: Integration of IT Service Management into Enterprise Architecture. Proceedings of the 2007 ACM Symposium of Applied Computing, Seoul, Korea, 2007, S. 1215-1219.
- [CI05] CITC-EASC: California Enterprise Architecture Framework, July 2005, Release 1.0 Final. California Information Technology Council (CITC) Enterprise Architecture and Standards Committee (EASC) Office of the State Chief Information Officer (SCIO), Sacramento, CA, 2005.
- [CI08] CIM: Common Information Model Standards. <http://www.dmtf.org/standards/cim/>. Abruf am 11.03.2008.
- [CI99] CIO-Council: Federal Enterprise Architecture Framework Version 1.1. <http://www.cio.gov/archive/fedarch1.pdf>, Abruf am 31.10.2006.
- [CJ07] Colville, R.; James, G.: Toolkit Decision Framework: EA Tools and CMDBs: Similar but Different for Some Time to Come. In: Gartner for IT leaders 23 August 2007 (2007).
- [EP00] Eriksson, H.-E.; Penker, M.: Business Modelling with UML – Business Patterns at Work. John Wiley & Sons, New York, 2000.
- [FA07] Fischer, R.; Aier, S.; Winter, R.: A Federated Approach to Enterprise Architecture Model Maintenance. In: Enterprise Modelling and Information Systems Architectures 2 (2007) 2, S. 14-22.
- [IE00] IEEE: Recommended Practice for Architectural Description of Software Intensive Systems (IEEE Std 1471-2000). New York, NY, 2000.
- [JL06] Jonkers, H. et al.: Enterprise architecture: Management tool and blueprint for the organisation. In: Information Systems Frontiers 8 (2006), S. 63-66.
- [Ke07] Keller, W.: IT-Unternehmensarchitektur: Von der Geschäftsstrategie zur optimalen IT-Unterstützung. dpunkt, Heidelberg, 2007.
- [KL05] Klint, P.; Lämmel, R.; Verhoef, C.: Towards an engineering discipline for grammarware. In: ACM Transactions on Software Engineering and Methodology (TOSEM) 14 (2005) 3, S. 331-380.
- [Kr05] Krupinski, A.: Unternehmens-IT für Banken: Kursbuch für IT-Management und IT-Architektur im Bankbereich. Vieweg, 2005.
- [KW07] Kurpjuweit, S.; Winter, R.: Viewpoint-based Meta Model Engineering. In: Proceedings, 2nd International Workshop on Enterprise Modelling and Information Systems Architectures, St. Goar/Rhine, 2007, S. 143-161.
- [La05] Lankhorst, M.: Enterprise Architecture at Work: Modelling, Communication and Analysis. Springer, Berlin et al., 2005.

- [LJ06] Lindström, L. et al.: A survey on CIO concerns-do enterprise architecture frameworks support them? In: Information Systems Frontiers 8 (2006) 2, S. 81-90.
- [Lu00] Luftman, J. N.: Assessing Business-IT Alignment Maturity. In: Communications of the Association for Information Systems (AIS) 4 (2000) 14, S. 1-50.
- [Lu05] Luftman, J. N.: Key Issues for IT Executives 2004. In: MISQ Executive 4 (2005) 2, S. 269-285.
- [MB05] Moser, C.; Bayer, F.: IT Architecture Management: A Framework for IT-Services. In: Proceedings, Moser.Bayer2005-ITArchitectureManagement.pdf 2005, S. 137--151.
- [Ni05] Niemann, K. D.: Von der Unternehmensarchitektur zur IT-Governance: Leitfaden für effizientes und effektives IT-Management. (Edition CIO) edition, Vieweg, 2005.
- [OM05] OMG: Unified Modeling Language: Superstructure. 2005.
- [OW07] Österle, H. et al.: Business Engineering – Core-Business-Metamodell. In: WISU – Das Wirtschaftsstudium 36 (2007) 2, S. 191-194.
- [Pa06] Pavlak, A.: Enterprise Architecture: Lessons from Classical Architecture. In: Journal of Enterprise Architecture 2 (2006) 2, S. 20-28.
- [PS05] Pereira, C. M.; Sousa, P.: Enterprise Architecture: Business and IT Alignment. In: Proceedings, 2005 ACM symposium on Applied computing (SAC 2005), Santa Fe, New Mexico, 2005, S. 1344-1345.
- [RT08] RTC. <http://www.rtc.ch/unternehmen/rtc-in-kuerze/>, Abruf am 11.03.2008.
- [SL06] Simonsson, M. et al.: Scenario-Based Evaluation of Enterprise Architecture – A Top-Down Approach for CIO Decision-Making. In: Proceedings, International Conference on Enterprise Information Systems 2006, S. 130-137.
- [SO02] SOX: The Sarbanes Oxley Act. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_bills&docid=f:h3763enr.tst.pdf, Abruf am 11.03.2008.
- [TC06] Tyler, D. F.; Cathcart, T. P.: A structured Method for Developing Agile Enterprise Architectures. In: Proceedings, International Conference on Agile Manufacturing (ICAM 2006), Norfolk, Virginia, 2006, S. 1-8.
- [To07] TheOpenGroup: The Open Group Architecture Framework TOGAF – 2007 Edition (Incorporating 8.1.1). Van Haren, Zaltbommel, 2007.
- [WB07] Winter, R. et al.: Analysis and Application Scenarios of Enterprise Architecture – An Exploratory Study. In: Journal of Enterprise Architecture 3 (2007) 3, S. 33-43.
- [WF07] Winter, R.; Fischer, R.: Essential Layers, Artifacts, and Dependencies of Enterprise Architecture. In: Journal of Enterprise Architecture 3 (2007) 2, S. 7-18.
- [Wo06] Wortmann, F.: Entwicklung einer Methode für die unternehmensweite Autorisierung. Dissertation, Institut für Wirtschaftsinformatik Universität St. Gallen 2006.
- [Za87] Zachman, J. A.: A Framework for Information Systems Architecture. In: IBM Systems Journal 26 (1987) 3, S. 276-292.

Integrationsmanagement mittels Quality Gates

Ein Erfahrungsbericht aus der Praxis

Tanja Snaikus*, Georg Disterer**

*OE Freigabemanagement
FinanzIT GmbH
Laatzener Str. 5
30539 Hannover
tanja.snaikus@finanzit.com

**Fakultät für Wirtschaft und Informatik
Fachhochschule Hannover
Ricklinger Stadtweg 120
D-30459 Hannover
georg.disterer@fh-hannover.de

Abstract: Beschrieben wird ein in der FinanzIT GmbH eingesetztes Verfahren zur Unterstützung des Integrationsmanagements, das auf Quality Gates basiert. Mit diesem Verfahren werden alle beteiligten Fachabteilungen des Unternehmens in die Qualitätskontrolle bei der Anwendungsentwicklung eingebunden. Damit wird eine sichere Integration von Anwendungen in eine komplexe und heterogene Systemlandschaft vorbereitet. Das erfolgreiche Durchschreiten der Quality Gates ist Voraussetzung für die Freigabe zur weiteren Entwicklung der Anwendungen und letztlich zu ihrem produktiven Einsatz.

1 Einleitung

Die FinanzIT GmbH (im folgenden Text: FinanzIT) ist mit rund 2.700 Mitarbeitern das IT-Systemhaus für Sparkassen, Landesbanken und Landesbausparkassen in elf Bundesländern. Mitarbeiter an den Standorten in Berlin, Hannover, Leipzig und Saarbrücken verbessern mit IT-Lösungen die Wettbewerbsfähigkeit von rund 170 Sparkasseninstituten und Banken. Die FinanzIT bietet ihren Kunden die Entwicklung und Integration von Anwendungssystemen, Hard- und Software-Beratung sowie IT-Services und betreibt leistungsfähige Rechenzentren. Die Systemlandschaft ist – wie bei IT-Dienstleistern dieser Größenordnung typisch – durch große Komplexität und Heterogenität gekennzeichnet, so dass die Integration neuer oder geänderter Anwendungssysteme anspruchsvoll ist.

Im Rahmen der Entwicklung und Inbetriebnahme von Anwendungssystemen werden bei der FinanzIT Quality Gates eingesetzt. Mit dieser Vorgehensweise wird der Anspruch

vertreten, ein flexibles und werkzeugunterstütztes Verfahren für das Integrationsmanagement zu schaffen und alle beteiligten Fachabteilungen und -gruppen (z. B. Netzwerkservices, Architektur, Sicherheit, Datenschutz) besser in den Prozess der Softwarebereitstellung und Integration einzubinden. Damit wird über die Unterstützung der technischen Integration betrieblicher Anwendungssysteme hinaus die umfassende Einbindung aller Stakeholder eines Entwicklungsprojekts angestrebt. Dabei sind funktionale und nicht-funktionale Anforderungen an das Entwicklungsergebnis sowie methodische Anforderungen an das Entwicklungsprojekt zu berücksichtigen, um eine Einbettung in die Geschäftsprozessarchitektur, die technische Architektur sowie das interne Prozessmodell des Unternehmens zu erreichen.

Im Folgenden wird beschrieben, wie das ursprüngliche Vorgehen zur Qualitätssicherung und -kontrolle schrittweise über eine begleitende Qualitätskontrolle zum heute eingesetzten Vorgehen mit Quality Gates weiterentwickelt wurde. Dabei werden die Methode, die Organisation sowie die Werkzeugunterstützung erläutert. Abschließend werden Ergebnisse und Erfahrungen der Einführung von Quality Gates aufgezeigt.

2 Ausgangslage

Bei der Anwendungsentwicklung und -inbetriebnahme werden in der FinanzIT die Qualitätssicherung und -kontrolle in einem zweistufigen Verfahren durchgeführt. Die Qualitätssicherung wird von den Entwicklungsprojekten verantwortet und schließt übliche Teststufen wie Modultest, Systemtest, Releasetest und Dokumententest ein.

Bei der Qualitätskontrolle wurde ursprünglich nach Abschluss der Entwicklungsarbeiten über die Freigabe der Software zum produktiven Einsatz entschieden.

Die Qualitätskontrolle wird vom Freigabekoordinator mit Stichprobentests durchgeführt, die folgenden Kategorien zuzuordnen sind:

- Funktionale Tests werden vom Freigabekoordinator auf einem Testsystem ausgeführt.
- Prüfungen der Einhaltung nichtfunktionaler Anforderungen sowie interner und externer Standards.
- Prüfungen bzgl. vom Unternehmen vorgegebener Regeln und Vorschriften durch den Freigabekoordinator stellen fest, ob die Prozesse zur Entwicklung und Qualitätssicherung einer Software eingehalten wurden.
- Prüfungen durch den Freigabekoordinator, ob die zuständigen Einheiten des Unternehmens eingebunden sind und ihre Zustimmungen – etwa in Form von Abnahmen – erteilt haben; etwa ob das Sicherheitskonzept abgenommen wurde.

Dieses wasserfallartige Vorgehen von Qualitätssicherung und Qualitätskontrolle zeigt u.a. zwei Probleme. Die Kriterien der Qualitätskontrolle sind nicht ausreichend transparent, da den Entwicklungsverantwortlichen nicht ausreichend klar ist, auf welchen Kriterien die Qualitätskontrolle basiert. Diese Kriterien werden erst deutlich, wenn die Quali-

tätskontrolle Abweichungen festgestellt hat. Viele der Abweichungen wären im Vorfeld zu vermeiden gewesen, wenn früher im Entwicklungsprojekt bekannt gewesen wäre, welche Kriterien der Qualitätskontrolle zugrunde liegen.

Zudem konzentrieren sich Aktivitäten der Qualitätskontrolle auf einen begrenzten Zeitraum kurz vor Inbetriebnahme der Software, in dem neben funktionalen Tests auch Dokumententests und Tests zur Einhaltung von Standards wie Sicherheit und Datenschutz durchzuführen sind. Daher führt eine nicht vorbehaltlose Abnahme oder ein Review, das Änderungen der Konzeption auslöst, ggf. zu einer Verschiebung der Inbetriebnahme.

Daher soll eine „entwicklungsbegleitende“ Qualitätskontrolle eine stärkere Verzahnung von Entwicklung/Qualitätssicherung und Qualitätskontrolle herbeiführen.

3 Entwicklungsbegleitende Qualitätskontrolle

Folgende Anforderungen wurden an die Weiterentwicklung der Vorgehensweise gestellt:

- Die Kriterien, auf der die Stichprobentests der Qualitätskontrolle basieren, sind den Entwicklungsprojekten frühzeitig bekannt zu geben. Damit können die Projektleitungen die Anforderungen der Qualitätskontrolle in die Planungen aufnehmen.
- Die Aktivitäten der Qualitätskontrolle werden soweit wie möglich begleitend zum Entwicklungsprozess durchgeführt, damit Abweichungen ggf. früher festgestellt werden. Damit löst die Behebung der Fehler geringere Kosten aus und drohende Auswirkungen auf den Termin der Inbetriebnahme können durch geeignete Maßnahmen des Projektmanagements besser abgefangen werden.
- Die grundsätzliche verfahrenstechnische Trennung zwischen Qualitätssicherung und Qualitätskontrolle soll aufrechterhalten bleiben, damit die Qualitätskontrolle weiterhin unabhängig und ohne Interessenskonflikte vom Entwicklungs-/Qualitätssicherungsprozess durchgeführt wird.
- Die in den Fachabteilungen etablierten und bewährten Vorgehensweisen und Prozesse sollten möglichst erhalten bleiben, um den Änderungsaufwand und -dauer bei Einführung einer entwicklungsbegleitenden Qualitätskontrolle zu beschränken.

Umgesetzt werden diese Anforderungen mithilfe einer Checkliste, die etwa drei Monate vor der Inbetriebnahme von der Leitung des Entwicklungsprojektes und dem Freigabekoordinator bearbeitet wird. In dieser Checkliste sind alle Ergebnistypen und zu erbringenden Leistungen enthalten, auf die sich die Prüfungen im Rahmen der Qualitätskontrolle des Freigabemanagements beziehen. Zu jeder Position der Liste wird vereinbart, ob sie für das aktuelle Entwicklungsvorhaben aus Sicht des Freigabekoordinators relevant ist oder nicht. Zu relevanten Positionen wird der Termin festgelegt, zu dem die jeweilige Anforderung erfüllt sein soll und damit für die Prüfung durch den Freigabekoordinator verfügbar ist.

Dann wird entwicklungsbegleitend vom Freigabekoordinator geprüft, ob die abgestimmten Positionen in der erforderlichen Qualität zum vereinbarten Termin zur Verfügung stehen. Damit werden Zwischenergebnisse wie Fachkonzept, DV-Konzept, Migrationskonzept u.a. feingranularer als vorher eingeplant und der Qualitätskontrolle zugeführt.

In der abschließenden Qualitätskontrolle werden die funktionalen Tests durchgeführt. Aufgrund der Ergebnisse aus der entwicklungsbegleitenden und der abschließenden Qualitätskontrolle entscheidet der Freigabekoordinator die Freigabe der Software zur Inbetriebnahme.

4 Weiterentwicklung der Qualitätskontrolle mit Quality Gates

Für eine weitere Verbesserung der entwicklungsbegleitenden Qualitätskontrolle wurden von den Fachabteilungen, dem Freigabemanagement sowie den Projektleitungen folgende Anforderungen gestellt:

- Die Fachabteilungen sollen unmittelbar in die Qualitätskontrolle einbezogen werden, d.h. die relevanten Kriterien zur Prüfungen sollen mit den jeweilig zuständigen Fachabteilungen abgestimmt werden.
- In die jeweiligen Prozesse der Fachabteilungen soll nicht eingegriffen werden. Die Fachabteilungen prüfen in eigener Verantwortung und nach ihren Verfahren und geben anschließend das Prüfergebnis in Form einer Statusmeldung an den Freigabekoordinator.
- Die Statusmeldungen werden zu einheitlich vorgegebenen Terminen (Quality Gates) final abgegeben. Das Meldeverfahren soll mit einem einheitlichen DV-Werkzeug unterstützt werden; dafür ist ein bereits im Unternehmen etabliertes Werkzeug zu nutzen.
- Das Verfahren soll für alle Beteiligten, insbesondere für Mitglieder der Entwicklungsprojekte, frühe Signale und Transparenz der Qualitätskontrolle erzeugen.

Zur Koordination der Statusmeldungen wurde auf ein im Innovationsmanagement etabliertes Ablaufmodell mit Quality Gates zurückgegriffen, das die Zusammenarbeit verschiedener Beteiligter an komplexen Prozessen wie etwa beim Management von Produktinnovationen, bei Produktanläufen und bei Produkteinführungen steuert. Der Einsatz von Quality Gates basiert auf der Annahme, eine hohe Qualität des zu entwickelnden Produkts sei nur mit einer hohen Prozessqualität bei der Entwicklung zu erreichen. Daher wird zur Sicherung der Prozessqualität eine dezidierte Prozesslenkung eingesetzt, um die Abstimmung und Synchronisation verschiedener Disziplinen mit einem strengen Ablaufmodell durchzusetzen. So sollen Fehler und Unklarheiten frühzeitig erkannt und behoben werden können. Die Zusammenarbeit und Abstimmung zwischen den Beteiligten wird mit Quality Gates Regeln unterworfen, die sicher stellen, dass zu festgelegten Zeitpunkten der Stand der Entwicklung und die Erfolgchancen der Fortentwicklung von allen Beteiligten klar und unmissverständlich kommuniziert wird.

Die Frühzeitigkeit wird dabei durch die Definition von Zeitpunkten (Terminen der Quality Gates) im Verlauf der Entwicklung sichergestellt, zu denen sich alle Beteiligten über die Qualität des Zwischenstands der Entwicklung abzustimmen haben. Die Klarheit und Unmissverständlichkeit der Abstimmung wird dadurch sichergestellt, dass alle Beteiligten ausdrücklich einer Fortentwicklung zustimmen müssen, indem sie mit einer Statusmeldung ihre Freigabe erteilen, andernfalls wird die Entwicklung gestoppt. Somit sind Quality Gates Koordinations- und Entscheidungspunkte, zu denen alle Beteiligten mit den Statusmeldungen dem Fortgang eines Entwicklungsprozesses ausdrücklich zustimmen bzw. widersprechen müssen. Dieser Druck zur positiven Stellungnahme oder ausdrücklichen Ablehnung soll Fehler und Unklarheiten frühzeitig und sicher aufdecken. Eine spätere Reklamation mit dem Hinweis auf mangelnde Beteiligung und Einbindung soll damit ausgeschlossen werden. Quality Gates dienen somit der Koordination verschiedener Beteiligter im Entwicklungsprozess; ein Entwicklungsprojekt hat ein Quality Gate nur erfolgreich durchschritten, wenn alle Statuslieferanten ihre Freigabe erteilt haben.

Mit Quality Gates werden alle Beteiligten eines Entwicklungsprozesses als Statuslieferanten fest eingebunden. Dabei fungieren als Statuslieferanten für funktionale Anforderungen Experten der jeweiligen Geschäftsprozesse, die zu unterstützen sind. Für nicht-funktionale Anforderungen primär Vertreter der zuständigen Fachabteilungen der Produktion sowie für methodische Anforderungen die Prozessverantwortlichen fachlicher Gruppen wie Qualitätsmanagement, Projektmanagement, Controlling, Risikomanagement. Damit ist eine enge Bindung aller Beteiligten an das Entwicklungsprojekt erreicht, da mit der Rolle als Statuslieferant das Recht auf ausreichende Information über das Entwicklungsprojekt, aber auch die Pflicht zur Abgabe von Statusmeldungen zu vorgegebenen Terminen verbunden ist.

Die entwicklungsbegleitende Qualitätskontrolle wird insoweit erweitert, dass nunmehr die Kriterien zur Prüfung und Freigabe einer Software nicht mehr nur zwischen den Entwicklungsprojekten und dem Freigabekoordinator, sondern auch mit den Statuslieferanten detailliert abgestimmt werden, die die Prüfungen der Kriterien durchführen.

Nach diesem Verfahren wird für jedes Projekt während des Projektstarts eine Anzahl von Quality Gates als Koordinations- und Entscheidungspunkte terminiert, und in einer Freigabecheckliste vermerkt, welche Prüfkriterien der Statuslieferanten welchen Quality Gates zugeordnet sind. Abbildung 1 zeigt beispielhaft für ein Projekt einen Ausschnitt aus der Freigabecheckliste und weist für einen Prüfbereich – beispielsweise „Netzwerkservices“ – einige Prüfkriterien sowie die Zuordnung zu den Quality Gates aus. Im Beispiel für diesen Bereich sind einige Prüfkriterien den Quality Gates „Vorbereitung“ und „Konzeption“ zugeordnet und der für die Prüfung und Statusmeldung Zuständige benannt. Unter „Relevanz“ werden Besonderheiten eines Entwicklungsprojektes abgebildet und begründet, etwa wenn spezielle Prüfungen wegen Spezifika des Projektes ganz entfallen können.

Prüfbereich	Inhalt	Relevanz	Quality Gate					Anspruchspartner
	Prüfthemen	Begründung für Abweichung / Bemerkung	Relevant für dieses Verhalten (Ja/Nein)	Vorbereitung	Konzeption	Entwicklungsende	Produktionsübergabe	Freigabe / Inbetriebnahme
Netzwerkkooperationen								
	Neuentwicklung: Es ist geklärt, ob die Anwendung Daten über das Netzwerk überträgt. Falls keine Netzwerkübertragung, sind alle weiteren Checkpunkte nicht relevant. Die zu erwartende Anwendungsnutzung / des Nutzerverhalten in Bezug auf die Tageszeit ist dargestellt		ja	x				Musterfirma, Max
	Es ist festgelegt, welche Kommunikationskanäle in welcher Form zu erwarten sind	keine Änderungen an den bisherigen Szenarien	nein		x			

Abbildung 1: Auszug aus der Freigabecheckliste (Beispiel)

In der generischen Beschreibung des Verfahrens sind bei der FinanzIT sieben Quality Gates vorgesehen, deren Inhalte an die Spezifika der einzelnen Entwicklungsprojekte angepasst werden. Diese Quality Gates sind wie folgt gekennzeichnet:

- Das Quality Gate „Vorbereitung“ wird am Ende der Erstellung eines Projektantrags – vor Entscheidung der Projektsteuerungsgremien über den Start des Projekts – durchgeführt. Mit diesem Quality Gate wird sichergestellt, dass dem Entscheidungsgremium nur validierte Informationen im Projektantrag, Fachkonzept und anderen Unterlagen vorgelegt werden.
- Das Quality Gate „Konzeption“ wird am Ende der Konzeptionsphase eines Entwicklungsprojektes durchgeführt. Damit soll gewährleistet werden, dass bereits in der Konzeptionsphase alle rechtlichen, organisatorischen, fachlichen und technischen Rahmenbedingungen berücksichtigt werden.
- Mit dem Quality Gate „Entwicklungsende“ soll gewährleistet werden, dass die Voraussetzungen für die Durchführung der abschließenden Qualitätskontrolle vollständig erfüllt sind und alle notwendigen Informationen für die Inbetriebnahme der Software vorliegen.
- Das Quality Gate „Produktionsübergabe“ wird zur Übergabe der Software von der Anwendungsentwicklung in die Produktion durchgeführt. Die Prüfungen anlässlich dieses Quality Gates stellen sicher, dass die Produktionsbereiche der FinanzIT ausreichend auf die Aufnahme des Pilot- oder Produktivbetriebs der Software vorbereitet sind.
- Die Quality Gates „Pilotierungsbeginn“ und „Pilotierungsende“ werden durchgeführt, wenn eine Pilotierungsphase vorgesehen ist. Die beiden Quality Gates stellen sicher, dass die erforderlichen organisatorischen und technischen Rahmenbedingungen für eine Pilotierung erfüllt sind und die Pilotierung ordnungsgemäß durchgeführt und ausgewertet werden kann.

- Das Quality Gate „Lieferfähigkeit“ gewährleistet, dass alle Voraussetzungen für den Regelbetrieb der Software erfüllt sind und somit die uneingeschränkte Lieferfähigkeit gegenüber den Kunden vorliegt.

Die Steuerung der Statusmeldungen für alle Quality Gates wird mit Hilfe eines Ticketverwaltungssystems durchgeführt, das die Verfolgung von Arbeitsaufträgen erlaubt und das Regelwerk zur Durchführung von Quality Gates unterstützt. Das Regelwerk umfasst Vorgaben u.a. zu Terminen, Zuständigkeiten, Abläufen, Kommunikationswegen, Entscheidungs- und Ausnahmeregelungen.

Für jedes Quality Gate eines Projektes wird zum Projektstart ein Ticket mit einem Fälligkeitstermin aufgesetzt. Die Tickets stehen für die Statuslieferanten als Prüfauftrag und als Pflicht zur Statusmeldung. Die Statuslieferanten melden die finalen Ergebnisse ihrer Prüfungen zu den Terminen der Quality Gates. Im Vorfeld eines Fälligkeitstermins sind die Verantwortlichen des Entwicklungsprojekts sowie die Statuslieferanten in der Pflicht, Abstimmungsgespräche bezüglich der Prüfkriterien zu führen sowie die erforderlichen Informationen auszutauschen. Die Statuslieferanten bearbeiten dann zu gegebener Zeit ihre Arbeitsaufträge durch Prüfung und entsprechende Meldungen, die durch Ampelsignale kommuniziert werden: Status „grün“ steht für „go“, d.h. alle Prüfungen verliefen erfolgreich und die Freigabe zur Fortsetzung der Entwicklung wird erteilt, „rot“ steht für „Stopp“, der Statuslieferant sieht ernste Gefahren für das Entwicklungsprojekt bzgl. Zeit/Kosten/Qualität und erteilt keine Freigabe; die Gründe für diese Beurteilung sind vom Statuslieferanten zu dokumentieren.

Wenn alle Statusmeldungen zu einem Quality Gate „grün“ signalisieren, wird der Gesamtstatus für dieses Quality Gate auf „grün“ gesetzt und das Gate ist erfolgreich absolviert. Andernfalls wird vom Freigabekoordinator eine Eskalation ausgelöst, um die zu Tage getretenen Probleme zu lösen und eine Entscheidung zum weiteren Fortgang des Entwicklungsprozesses herbeizuführen. Die Eskalation beginnt i.d.R. mit einem vom Freigabekoordinator moderierten Gespräch zwischen der Leitung des Entwicklungsprojekts und den Statuslieferanten, die Bedenken haben. Hinzugezogen werden die jeweiligen Vorgesetzten. Wird in diesem Kreis keine Einigkeit über den Status erzielt, wird die Entscheidung zur Fortführung, wesentlichen Änderung oder Beendigung des Projekts der Geschäftsführung vorgelegt.

Zur Ticketverwaltung wird ein Werkzeug eingesetzt, das bei der FinanzIT etabliert ist und für das Incident-, Problem- und Change-Management genutzt wird. Damit ist die überwiegende Anzahl der Mitarbeiter mit den wesentlichen Funktionen des Werkzeugs vertraut. Zur Unterstützung der Tickets für Quality Gates wurde ein spezielles Modul „Freigabemanagement“ eingerichtet, mit dem neue Tickets erstellt und vorhandene Tickets gesucht und bearbeitet werden können.

Der Ablauf des Verfahrens ist als Prozess modelliert; Abbildung 2 zeigt eine schematische Darstellung.

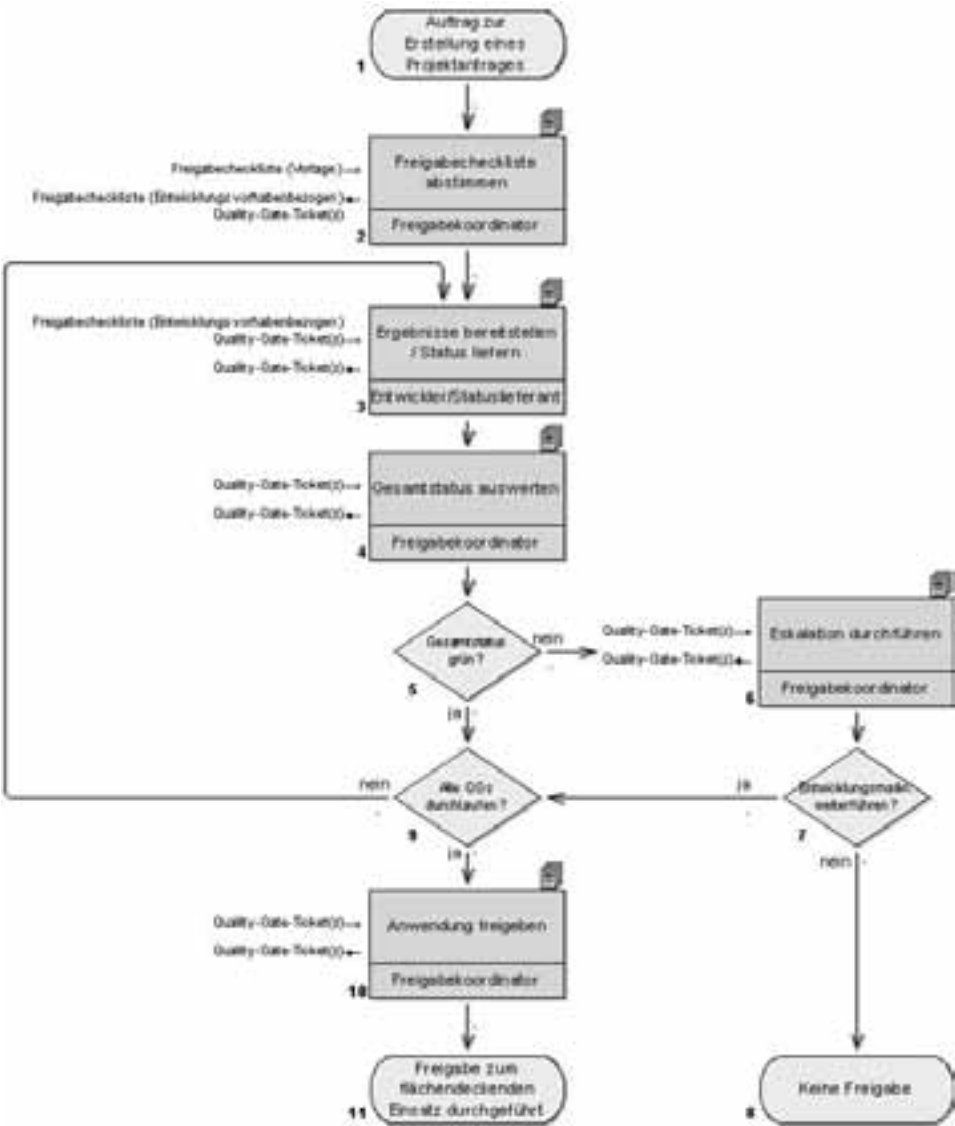


Abbildung 2: Schematische Darstellung des Verfahrens mit Quality Gates

Abbildung 3 zeigt beispielhaft ein Ticket, das zu dem Quality Gate „Vorbereitung“ eines Projektes „Testprojekt“ eingerichtet wurde. Für die grünen Einzelprüfungen liegen bereits positive Statusmeldungen der zuständigen Lieferanten vor. Bei den gelben Einzelprüfungen haben die Statuslieferanten schon vor Ablauf der Bearbeitungsfrist Bedenken gemeldet, um diese ggf. noch auflösen zu können. Der rote Eintrag steht für eine Meldung, mit der ein Statuslieferant eine Freigabe verweigert hat. Jedem Prüfbereich ist ein für die Statuslieferung verantwortlicher Mitarbeiter namentlich zugeordnet.

Mit dieser Übersicht können der Freigabekoordinator sowie die Projektleitungen und andere Beteiligte jederzeit einen Überblick zu jedem Projekt erhalten. Weitere Details zu den Einzelprüfungen werden über das Öffnen einzelner Zeileneinträge der Liste aufgerufen.

Ticket Nr.
555

Titel
Testprojekt / Quality Gate „Vorbereitung“

Ticket Ersteller
Otto Normalverbraucher

Freigabekoordinator
Erika Mustermann

Termin
24.04.2008

Gesamtstatus
rot

Prüfbereich	Statuslieferant	Status
Prüfbereich 1	Statuslieferant 1	grün
Prüfbereich 2	Statuslieferant 2	gelb
Beispielprüfbereich „Netzwerkservices“	Max. Mustermann	rot
Prüfbereich 4	Statuslieferant 4	grün
Prüfbereich 5	Statuslieferant 4	gelb

Zugeordnete Changes	Verantwortlicher Mitarbeiter
88888 Aktivität 1	Mitarbeiter 1
88889 Aktivität 2	Mitarbeiter 2

Historie anzeigen

Abbildung 3: Ticket zum Quality Gate „Vorbereitung“ für ein Testprojekt (Beispiel)

Weiterhin können alle Changes eingesehen werden, die zum selben Entwicklungsprojekt vorliegen. Mit diesen Changes werden mit dem Entwicklungsprojekt in Verbindung stehende Änderungen an der Produktionsumgebung und der Systemsoftware dokumentiert. Für jede identifizierte Änderung wird ein Change erfasst, in dem diese nachvollziehbar dokumentiert ist. Damit wird transparent, welche Changes zu einem Entwicklungsprojekt vorliegen und welche Quality Gates für dieses Entwicklungsprojekt vorgesehen bzw. schon absolviert sind. Auch die Änderungshistorie zu einem Ticket kann eingesehen werden.

5 Einführung

Die Einführung der Vorgehensweise mit Quality Gates für Großprojekte geschah im Zeitraum Mitte 2006 bis Mitte 2007. Die Konzeption und Umsetzung des Qualitätskontrollverfahrens mit Quality Gates wurde von vergleichsweise wenigen Mitarbeiter/innen durchgeführt. In dieser verfügbaren Zeit war das Verfahren zu entwickeln und zu beschreiben sowie Freigabekoordinatoren, Projektleiter und Statuslieferanten in die Vorge-

hensweise einzuführen. Die Mitarbeiter wurden dazu vorab über das Vorgehen und damit verbundene Neuerungen informiert, weitere Informationen wurden begleitend zur Einführung gegeben. Im Zuge der erstmaligen Einführung des Verfahrens bei den ersten Projekten bekamen die Mitarbeiter eine Einweisung sowie Unterlagen zum Selbststudium.

Zusätzlich waren folgende Detailarbeiten zu bewältigen:

- Eine generische Freigabecheckliste musste entwickelt werden, die allen neuen Projekten als Grundlage dienen soll, um daraus die projektspezifisch vorzusehenden Quality Gates abzuleiten. Eine erste Zusammenstellung ergab dabei 36 verschiedene Prüfbereiche mit 1.450 einzelnen Prüfkriterien. Nach Konsolidierung umfasst die Liste nunmehr 30 Prüfbereiche und insgesamt 250 einzelne Prüfkriterien, die projektspezifisch auf sieben Quality Gates zu verteilen sind.
- Die Werkzeugunterstützung war zu planen und umzusetzen. Dabei war auch die Verbindung zum Change-Verfahren zu realisieren. Eine erste Version des Moduls zum Freigabemanagement stand im Dezember 2006 zur Verfügung.
- Die Einführung des Verfahrens war durch den Prozessverantwortlichen für das Verfahren intensiv zu begleiten; dabei wurden aus der praktischen Umsetzung notwendig erscheinende Anpassungen an dem Verfahren vorzunehmen.
- Für neue Großprojekte war die Qualitätskontrolle mit Quality Gates initial aufzusetzen und durchzuführen.
- Für laufende Großprojekte war gemeinsam mit den Projektleitungen zu bewerten, ob eine Einführung der Quality Gates für die verbleibende Projektlaufzeit sinnvoll ist. Bei Umstellungen auf das Verfahren mit Quality Gates waren die Inhalte der für die verbleibende Projektlaufzeit vorzusehenden Quality Gates festzulegen und die Quality Gates durchzuführen.

Anfang des Jahres 2007 wurden Quality Gates für die ersten Großprojekte aufgesetzt und die Prüfungen zur Qualitätskontrolle im Rahmen des ersten Gates „Vorbereitung“ durchgeführt. Im ersten Halbjahr 2007 waren das Verfahren und die Werkzeugunterstützung weiter zu entwickeln und zu verfeinern und die ersten Erfahrungen aus der Nutzung von Quality Gates einzuarbeiten. Dabei konnten Fragen, Probleme und Änderungsbegehren der beteiligten Mitarbeiter direkt und schnell identifiziert und in die Weiterentwicklung aufgenommen werden.

Ab Mitte des Jahres 2007 wurde das Verfahren für mittlere und kleine Projekte angepasst. Dafür wurde die Anzahl maximal für ein Projekt vorzusehender Quality Gates auf drei begrenzt. Auch einige der Prüfkriterien konnten aus der Freigabecheckliste entfernt werden, da sie nur für Großprojekte relevant sind.

6 Ergebnis und Erfahrungen

Das Verfahren der Qualitätskontrolle mit Quality Gates ist bei der FinanzIT für große Projekte ab Anfang 2007, für mittlere und kleine Projekte ab Herbst 2007 flächendeckend im Einsatz. Dabei sind bisher 650 Quality Gates definiert worden, von denen 490 bereits durchlaufen wurden (Mitte 2008); bei den restlichen Gates stehen die Fälligkeitstermine noch an. Dabei entfallen ca. 18% der Gates auf Großprojekte, die restlichen 82% auf mittelgroße und kleine Projekte.

Folgende Beobachtungen aus dem Einsatz des Verfahrens spiegeln die Erfahrungen mit Quality Gates wider:

- Die Verantwortlichen für Entwicklungsprojekte berücksichtigen die Belange der Fachabteilungen stärker, da deren Einbindung in den Entwicklungsprozess durch die Rolle als Statuslieferanten verstärkt wurde.
- Die frühere und intensivere Einbindung der Statuslieferanten erzeugt eine deutlich höhere Transparenz über den Status von Entwicklungsprojekten bzgl. der Qualitätskontrolle.
- Die Anzahl notwendiger Eskalationen wird zum Ende von Entwicklungsprojekten geringer, da auftretende Probleme viel früher erkannt und deren Ursachen früher beseitigt werden.
- Die Fachabteilungen erkennen das Verfahren mit Quality Gates als Unterstützung für die Durchsetzung ihrer Belange an. Manche Fachabteilungen haben aktiv ihre Aufnahme in das Verfahren als Statuslieferanten angestrebt.
- Die Qualität der Ergebnistypen steigt, zunehmend werden Ergebnistypen auf Anrieb verfahrenskonform angefertigt und vorgelegt.

Die Einführung des Verfahrens mit Quality Gates wurde wesentlich durch folgend beschriebene Merkmale bestimmt.

Flexibilität: Die Inhalte und Anzahl der Quality Gates werden für jedes Entwicklungsprojekt angepasst. Damit ist sichergestellt, dass ein standardisiertes Verfahren passend für jedes Entwicklungsvorhaben eingesetzt wird.

Eigenverantwortung der beteiligten Fachbereiche (Statuslieferanten): Das Verfahren gibt den Fachabteilungen einen organisatorischen Rahmen für deren eigene Aktivitäten der Qualitätskontrolle vor. Damit wird ein Standardverfahren bereitgestellt, in das die Fachbereiche und Fachgruppen ihre spezifischen Prüfungen zur Ermittlung von Statusmeldungen einbringen können. Die Prüfungen und deren Durchführung verbleiben vollständig in der Verantwortung der Fachbereiche und -gruppen.

Kompetenz für das Freigabemanagement: Das Verfahren mit Quality Gates wird von der Abteilung Freigabemanagement verantwortet, die auch die Freigabeentscheidung zur Inbetriebnahme trifft. Für den erfolgreichen und schnellen Einsatz des Verfahrens war es wichtig, dass diese Abteilung mit den erforderlichen Kompetenzen und Ressourcen ausgestattet ist. Da die endgültige Entscheidung zur Inbetriebnahme einer Software vom Freigabemanagement zu treffen ist, waren die Entwicklungs- und Fachbereiche ausreichend motiviert, eine schnelle Einführung und Umsetzung der Quality Gates zu unterstützen.

Werkzeugunterstützung für das Verfahren: Eingesetzt wird ein im Unternehmen bewährtes Werkzeug, das eine Schnittstelle zum Change Management besitzt. Damit sind die Zusammenhänge zwischen Entwicklungsprojekten, Freigaben und Changes jederzeit transparent. Das Werkzeug kann an neue Anforderungen angepasst werden, so sind etwa Änderungen der standardmäßig vorgegebenen Prüfkriterien schnell durch administrative Eingriffe an den Parametern der Software vorzunehmen.

7 Fazit

Von der überwiegenden Zahl der Beteiligten ist die Qualitätskontrolle mit Quality Gates als positive Weiterentwicklung der bestehenden Vorgehensweisen aufgenommen und anerkannt worden.

Nennenswerte Probleme und Schwierigkeiten bei der Einführung des Verfahrens sind trotz der hohen Anzahl beteiligter Fachabteilungen und -gruppen nicht aufgetreten. Änderungen an der Freigabecheckliste und bei den Statuslieferanten können einfach berücksichtigt werden.

Mit dem Verfahren werden alle beteiligten Fachabteilungen des Unternehmens in die Qualitätskontrolle bei der Anwendungsentwicklung eingebunden, um die Integration der Anwendungen in eine komplexe und heterogene Systemlandschaft zu gewährleisten. Durch das flexible und werkzeugunterstützte Verfahren für das Integrationsmanagement werden alle beteiligten Fachabteilungen des Unternehmens besser in den Prozess der Softwarebereitstellung und Integration eingebunden.

Data Migration Project Management and Standard Software – Experiences in Avaloq Implementation Projects

Klaus Haller


COMIT AG
Pflanzschulstr. 7
CH-8004 Zürich
klaus.haller@comit.ch

Abstract: When a bank implements a new core banking platform, it must simultaneously migrate its data from the old platform into the new one. We identified three main dimensions for data migration. They make up the corners in our conceptual data migration project management framework: the *data migration triangle*. The first dimension demands scripts as deliverables, which actually transform and migrate the data. The second, and too often neglected, dimension is the delivery of data. It requires that the scripts run together reliably and on time in a complex environment with many dependencies. Dimension three addresses quality assurance. We describe methods for checking whether all necessary data is migrated and whether the data is migrated correctly. In contrast to previous work, we focus on technical issues from a project manager’s perspective. We explain *what* data migration project managers have to manage and *why*.

1 Motivation

For some years, a wave has flooded the banking IT landscape in Switzerland. Many banks have replaced their old banking applications with new standard software. They modernized their application landscape and reduced the total number of applications and interfaces. Thereby, the banks cut costs and eased out-sourcing. Avaloq [Ava] and Finnova [Fin] are the two dominant players on the Swiss market [Gab07, Min07]. But implementing one of the two is only the first challenge; the second is the data migration¹. During this transition, the banks’ CIOs and customers share the same fears: The new standard software platform replaced the old platform; the old one is shut down. But later, customers notice that the migrated data is incomplete or corrupt. For example, customers’ balances might be incorrect or account statements are mailed to customers who have asked the bank not to mail their statements.

¹ To prevent confusion: *Data migration* means migrating only data out of one schema to a new schema which can be structured completely differently. This is the case when a new application (e.g. Avaloq or Finnova) replaces an old one. Aim of *database migration* project is to change the database version or vendor, e.g. from Oracle 10g to Oracle 11g. Schema, triggers, date etc. should remain unchanged, if it is possible, because applications using the database should remain untouched.



Level	Examples
Project Management	Critical Path Method, Expected Return of Investment of Projects
IT Project Management	Rational Unified Process
Data Migration Management & Controlling	Key Performance Indicators, Butterfly Approach (Phases)
„Real Life“ IT View	Deployment
Data Migration Tools & Architecture	SAP solution, Butterfly Approach (Architecture)
Lab View	Implementation in PL/SQL

Figure 1: Data Migration Methodology Levels

The key for a successful data migration project is a superior team and a convincing methodology. The methodology in implementation projects is not monolithic but rather like a puzzle. Pieces of the puzzle come from the bank, the implementation partner, the project manager, and some can be chosen by the migration team. This paper offers a new piece of the puzzle: a cohesive lense for data migration project managers to look on their data migration projects, considering not only project organization and specification and implementation, but also the complexity of “real life” IT environments.

Figure 1 provides an overview about the different levels for methodologies relevant for data migration projects. The top level, the project management level, addresses IT and non-IT projects. A good example is the usage of the critical path method or tools like Microsoft Project in companies as well as the return of investment the projects are expected to generate. The second level from the top concentrates on IT specific project management issues like the software development methods (e. g. RUP [Kru04] or Extreme Programming [BA04]). One level down, we reach the management and controlling level for data migration projects. Quite some work has been published in this area. Morris [Mor06] provides an excellent introduction into data migration project organization and project management (though he also addresses underlying technological questions). Shorter articles like [BM04, Hud98] address some of the basic problems and pit falls you might encounter in your first data migration project. The butterfly approach [WLB97] contributes to this level with a high level phase model. Bisbal et al. [BLW97] provide a general discussion of migrating legacy information systems, i. e. the authors discuss more than the data migration aspect. Focusing more on data migration for standard software, [Hal08] provides a more in-depth discussion of the different phases. Also, this paper belongs to the “Data Migration Management & Controlling” level. Based on our experience in different Avaloq implementation projects, our paper describes the tasks of the lower levels with the aim of making their progress measurable.

We name the next level down the “real life” IT view. It includes everything else besides specification and implementation of the migration. Typical topics include integration problems between the data migration sources and all information systems as well as strategies for deploying data or the standard software on servers.

Data migration projects have specific tasks for which they use specific tools (e.g. ETL tools). Furthermore, they rely on a specific architecture. Important previous works contributing to this level include a description of SAP's data migration tools [WG04], a specific language [CG04], a data migration tool [DMS], or the architecture of the butterfly approach [WLB97]. [Hal08] provides a detailed architecture which also comes with suggestions for organizing the scripts. Finally, there is the "lab view" on data migration. The level focuses completely on the narrow aspect of specifying the migration and implementing the specification in PL/SQL scripts².

We structure our discussion of the different dimensions of data migration projects by focusing on a management view. Section 2 presents a big picture of the system architecture with specific focus on components and dependencies. Section 3 gives a quick overview on standard software implementation project organization. The data migration triangle, the core illustration and explanation concept, is the subject of Section 4. The triangle provides a high level overview of the three core data migration project dimensions. We elaborate the three dimensions in more detail in the following three sections: Section 5 concentrates on the transformation from the old to the new database schema (level "Lab View"). Section 6 discusses delivery issues, mainly integration and dealing with large amounts of data ("Real Life" IT view). Quality assurance is the topic of Section 7. It belongs conceptually to the data migration and controlling level, but an implementation would belong to the "lab view." It provides a risk model for the technical implementation and how the risks can be tracked. We conclude our paper with a short summary (Section 8).

2 System Architecture: Components and Dependencies

Standard software has three core components: a data model including domain value tables, operational data, and workflows (Figure). The *data model* defines the object types managed by the system. In contrast to software development projects, standard software typically restricts changes of the data model. Avaloq, for example, provides the object type "business partner." It must be used for customers, counter parties, etc. Avaloq does not allow adding a new object type "students." However, the data model gives the freedom to introduce a new attribute for "business partners" with values "student," "retired," etc. to model this semantic.

"Student" and "retired" are typical examples for domain value tables. They define a set of values an attribute can take. Another example for domain value tables are bank specific names for account types like "Super Savings Account" or "Senior Residents Checking Account".

Data models offer a data- and IT-centric view of the standard software. *Workflows* implementing business processes represent the business perspective. They specify how users, customers, back office employees, and support staff use the software. Depending

² Throughout the paper, we use the term script. Certainly, it is possible to use ETL-tools instead of writing scripts.

on the philosophy of the standard software, workflows can be defined freely, partially, or are completely unchangeable.

Rarely, banks buy standard software for completely new purposes. In most cases, they have been successfully operating for many years. So the bank has already valuable and important data. The old *data* must be migrated into the new system. The data will be complemented by new operational data as soon as the new software is operational.

There are dependencies between the migration of old data, workflows, and the data model. The data model is the foundation for the migration of old data and for the storage and processing of new data (Figure 2 ❶). The data model also influences the workflows (Figure 2 ❷). Workflows manipulate data and depend on data. If the data model does not allow storing a picture of the customer’s passport, the execution of a workflow cannot demand such a picture to be stored. Workflows can also put restrictions on data entered by users (Figure 2 ❸) e.g. by only accepting customers whose nationality is known. Depending on the standard software migration philosophy, old data is written directly into database tables or it is migrated using the same workflows as for manual data entry. In the latter case, there is a dependency between workflows and data migration (Figure 2 ❹). Figure 2 contains two more components. First, there is an underlying basic infrastructure with hardware, network, databases, and the actual standard software kernel. They are easily overlooked, but our experience shows their major negative impact on the overall project performance if not monitored and managed properly. Furthermore, there is a deployment infrastructure for storing, loading, and copying system snapshots, e.g. for distributing snapshots to different servers. A system snapshot consists of the data model, data, and workflows. Also the standard software kernel or even the complete database might be included. The snapshots are used for rerunning a migration from a certain point (e.g. with migrated customers to test the script for the customers’ account script over and over again) or for analyzing problems on different servers. It has not been our experience alone, but others have also observed that there is a huge impact on the development cycle whether it takes ten minutes or ten hours to reproduce a certain situation.

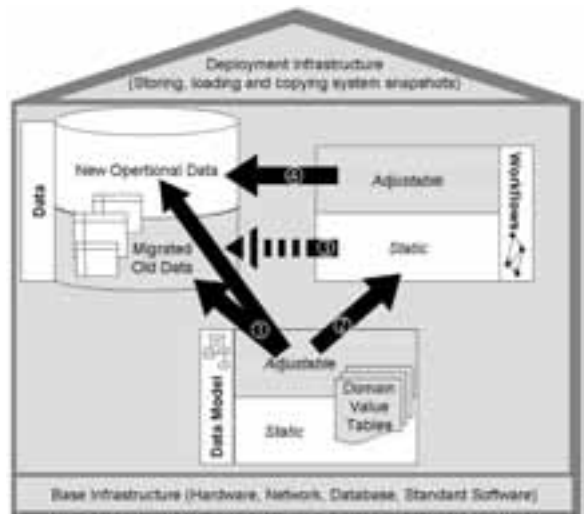


Figure 2: System Architecture with Components and Dependencies

3 Project Organization

We present a simplified version of the project organization approach of COMIT’s Lean Stream Implementation Method [LS07]. We distinguish four teams: a migration team, a customization team, an infrastructure team, and a support team (Figure 3). The *migration team* analyzes the data in the old system, specifies and implements the data mapping for transforming the old data to fulfil the demands of the new system, e.g. implemented in PL/SQL-scripts. Secondly, the migration team ensures the delivery. Delivery means orchestrating script execution and manual data entries. It is a challenging task because even small banks require hundreds of scripts for the migration. The migration team uses specialized migration tools which they manage themselves. Our experience shows that the highly specialized, in-depth knowledge of the migration team is urgently needed, but only until the going live stage of the project. Consequently, our project organization guarantees that the migration team can move to the next project to start over again directly after going live.

The *customization team* is responsible for the workflow implementation and adoption and the domain value tables. Team members usually have a sound knowledge of the business processes and of the standard software in their particular area. The project organization reflects this with a topic-oriented organization like accounting, CRM, etc. in the workflow team. Object modelling, highly important for software development, is a difficult topic in standard software implementation projects. The standard software comes with a predefined object model usually only requiring and allowing minor modifications. However, each data model change might require the migration team to reimplement already developed transformations. Thus, the migration team is highly interested in a stable and consistent object model. Consequently, it gets the responsibility for the object model, but it must collect continuously the requirements from workflow team members.



Figure 3: Project Organization

The *infrastructure team* configures and deploys the standard software, implements interfaces with satellite systems like trading, risk, or anti-money-laundering systems, and takes care of basic services such as hardware and network. If tasks like server or database administration are outsourced, the infrastructure team manages the supplier relationships.

The *support team* has two responsibilities. One is at least partially needed after the migration, namely to coordinate the testing with test case definitions, etc. This is also needed for later releases. Secondly, the support team is responsible for the initial user training. Even small banks have hundreds of employees to be trained.

4 The Data Migration Triangle

Projects have one thing in common: there are typically more things to be done than staff and time to do them. So it is important to set priorities. Setting priorities requires knowing the tasks. We group all data migration tasks into three dimensions, which are corners of our *migration triangle* (Figure 4). In this chapter, we give a brief overview, whereas we elaborate each task in more details in the following chapters.



Figure 4: Data Migration Triangle: Dimensions and Tasks

The first corner of the migration triangle represents the dimension called *mapping*. Mapping means defining and implementing the transformation³ of business objects. A business object is a meaningful object from a user’s perspective such as an account, a customer, or a customer’s address. The transformation ensures that the data fits into the target platform database schema. The dimension *mapping* has three tasks: the identification of the business objects to be migrated (task *business object identification*), the implementation of transformations for all business objects (task *business object complete-*

³ We use the term *transformation* for ease of speech in this paper in an inclusive way, meaning that it also comprises the extraction from the old and loading into the new system.

ness), and migrating all details of the different business objects (task *attribute mapping completeness*).

The second corner of the data migration triangle, *delivery*, includes correctly executing the different transformations and coordinating them with manual data entries (if not all of the data is migrated automatically). The delivery dimension consists of three tasks: migration-migration-integration (MIG/MIG-integration), migration-customization-integration (MIG/CUS-integration), and data set completeness. The task *MIG/MIG-integration* ensures that the different mapping implementations of the migration team run together. The task *MIG/CUS-integration* takes care that the workflows, domain value tables, and the data migration fit together. The task *data set completeness* tests carefully whether the migration runs in the estimated time for large data sets.

Quality assurance is the third dimension. Test cases must be defined and tested (task *test cases/testing*) for ensuring a semantically correct migration result. The task *reconciliation* guarantees that the migrated data is not only correct, but that also all needed data is migrated. If testing or reconciliation detects failures, they have to be solved earlier or later (task *failure reduction*).

The data migration triangle as a conceptual framework enables project managers to express and illustrate priorities in data migration projects and how priorities change during the project’s life-cycle. We illustrate this for the three typical situations we identified in our projects: the initial phase, the development phase, and the pre-going live phase. (Figure 5)

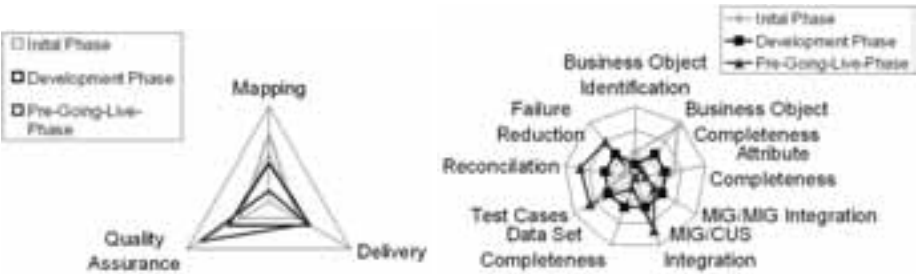


Figure 5: Data Migration Triangle and Extended Data Migration Radar View

Because there are dependencies between business objects, the migration team migrates a basic version of as many business objects as possible in the *initial phase*. Avaloq, for example, allows migrating a bank account only if the customer has been migrated before. Migrating an account balance is only possible if the account exists. Generally speaking, business objects form chains or, more accurately directed, acyclic dependency graphs. The “early” business objects must be migrated early in the project, even if they are not completely correct, such that work on the post-ordered business objects can start.

During the *development phase* all business object types should be looked at in depth, i.e. all attributes have to be considered, and the quality of the migration should be more or less good enough for going live.

A few weeks before going live, the *pre-going live phase* starts. The migration team must work on ensuring that (nearly) 100% of the objects are migrated automatically. It is especially important that the migration team monitors the data model, all domain value tables, and every workflow change of the customization team to prevent the going live migration from failing due to incompatibilities.

The priorities for the different phases are reflected in the data migration triangle and an extended radar view in Figure 5:

- To reach the goal of migrating as many business objects as possible, priority is given to business object completeness and MIG/MIG integration during the initial phase. Furthermore, test cases should be defined, though they cannot be tested extensively.
- During the development phase, the different subtasks are equally important, if no specific problems arise. The business objects' identification should be nearly completed and require only very limited affords.
- During the pre-going live phase, emphasis is given to a last round of failure reduction, so that not too many manual migration or corrections are required. Furthermore, the integration of the customization and the migration team is important. Systems like Avaloq allow changes of the data model, domain data tables, and migration relevant workflows till the last moment. So the migration team has to run frequent tests of their migration routines to prevent last time changes from causing major problems when going live under time pressure. Certainly, additional test cases are important and must be considered, including retesting previously failed ones.

5 The Mapping Task

The most obvious results of the migration team's work are executable scripts for extracting data from the old system, transforming it, and loading it into the new standard software. In a first step, the migration team identifies what shall be migrated (task *business object identification*). The team compiles a list with all business objects to be migrated like the first column of the table in Figure 6. Compiling this list is far from trivial. Theoretically, the migration team can start with an empty list and collect the business objects by interviewing the workflow team, examining the GUI and the current output, etc. However, there is a high risk of forgetting important business objects. We prefer to start with a copy of a previous project and to adopt the copy to the specific needs of the actual bank. The software vendor or experienced consulting companies should provide such a list for their customers.

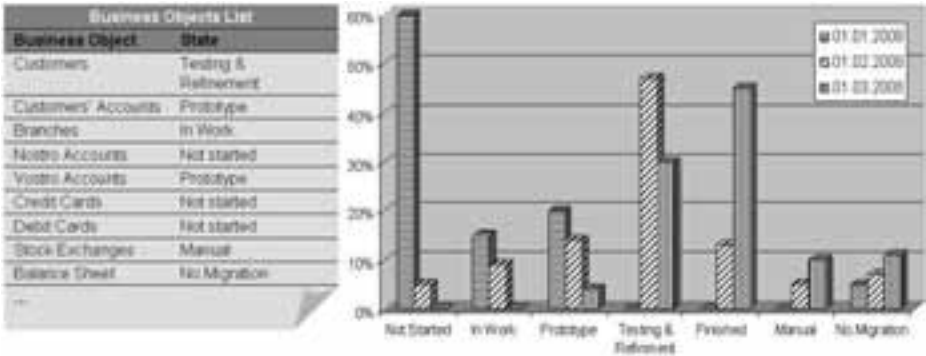


Figure 6: Business Object List (left) and 1

When the migration team has identified at least the most important business objects, the specification of the transformation starts.⁴ Our experience shows that a simple mapping sheet is sufficient (Figure 7). Starting points are the different windows of a GUI prompting the different data items of a business object, which we assume to be the most important information to be migrated. The migration team fills out a separate sheet (e.g. in Excel). Each attribute in a window is a row on the sheet. A row stores the information of the attribute's name in the window and the table and attribute name in the standard software. The table might be a table of a migration API or directly the internal table. Furthermore, the table and attribute name of the old system is needed. The migration type states whether the attribute is migrated normally ("migrate") or whether the information is derived automatically by the standard software ("derived" like the QI information which depends on the nationality etc.). It is also possible that no migration is needed ("no migration") or that due to the small amount of data, the data is manually entered ("manual"). If data comes from more than one table of the old system, the join condition between the different tables should be added. After the transformation is specified, it is implemented using SQL, PL/SQL, or an ETL-tool. When a migration run is finished, tests are performed and reconciliation tasks can start.

From a project manager's perspective, it is important to measure the progress of all three tasks. *Business object identification* progress corresponds to the stability of the list of the business objects (Figure 7, left). The key performance indicator is the number of newly added business objects, e.g. per week or month. The lesser new business objects are added, the more stable the list is, and the more likely it is already complete.

⁴ To our best knowledge, no well-established graphical specification techniques for data transformation exist (like we know e.g. UML or ER diagrams for data models), though certainly for the technical implementation much research in the context of schema mapping languages was done or is provided by existing ETL tools like [Pow].

Business object completeness reflects how many of the identified business objects are already migrated. Our experience shows that instead of a simple “done/to be done” model, a more complex model is appropriate: “not started” (nobody has done anything), “in work” (work has started but no result yet), “prototype” (first version successfully migrated), “testing and refinement” (it works in principle but not every attribute has the right value), “finished” (accepted by the customer). The progress is measured per business object. The overall situation is summed up in a statistic (Figure 7, right) for reporting purposes.

Attribute mapping completeness addresses whether all details of a business’ objects are considered. For example, a customer can be modeled in the beginning only by her name.



Figure 7: GUI View (left) and Specification Sheet (right)

So the business object exists which is helpful for business objects depending on the existents of addresses. However, all attributes of the addresses must be migrated in the end. Measuring the attribute mapping completeness is not really easy (and counting attributes in the old and new system usually only adds bureaucratic overhead). However, the already discussed statistics for business object completeness also contain this information (Figure 7, right).

6 Migration Delivery

When the migration team members have specified and implemented scripts each for themselves, the next two challenges are first, keeping the scripts compatible with the rest of the environment and the other migration scripts, and second, ensuring they run in a sensible time frame. We start our discussion with the latter challenge, the task of *data set completeness*. The underlying problem consists of two contradicting demands. The migration team needs short development cycles. If migrating the complete data of a bank needs several days, the migrated data is restricted for testing purposes to a few representative branches to reduce the migration time and thereby the development cycles. Secondly and contrary, the final migration duration often has an upper limit, because e.g. a bank cannot stop its operation for a week. So the migration team must identify and optimize long running scripts by tests migrations with the complete data set. Especially problematic are transformations with non-linear complexity. If migrating 1,000 customers needs one hour, migrating 10,000 customers might need 10 hours for a complexity of $O(n)$, but 100 hours for a complexity of $O(n^2)$. Also memory, network, or disk problems might suddenly occur. Thus, it is mandatory to run regular tests. Valuable indicators for the progress of this task are (i) frequency of uploads with the whole data set, (ii) the

needed time for completion for migrating all date with all scripts, and (iii) the top-n scripts regarding execution time (Figure 8).

Upload Date	Duration	Top-3 Scripts	
15.1.2008	64h	1. Customers	10h 25min
		2. Customer Accounts	9h 37min
		3. Branches	8h 12min
15.2.2008	50h 10min	1. Customer Accounts	9h 58min
		2. Customers	9h 4min
		3. Debit Cards	9h 27min

Figure 8: Indicators for Task Data Set C 1

The second aspect, *integration* respectively integration tests, can be further structured with respect to the involved teams. There is a customization/customization (CUS/CUS)-integration for testing, e.g. whether the workflow for cashiers for a high withdrawal in a foreign currency works properly with the involved forex dealer’s workflow. Frequently, the customization team builds a customization build. It is a consistent set of workflows and domain value tables of all customization team members. The customization build is then rolled out to all customization development servers and the test and release management cycle starts over again (Figure 9).

Of particular interest for us are integration issues involving the migration team. Similar to the customization release management cycle, the migration team performs integration tests by executing the scripts and thereby making a migration build (migration/migration-integration or, short, *MIG/MIG integration*). So they continuously check, for example, whether their scripts fit together with respect to primary/foreign key relationships. The migration release is deployed on the migration development servers. Usually, some migration development servers get an intermediate state after some, but not all scripts have been executed. Developers working e.g. on the migration of customer accounts want a server state where the customers are already migrated, but the customer accounts scripts have not been executed. So they can test improved versions of their scripts.

MIG/MIG-integration and CUS/CUS-integration are two separate release management cycles. Additionally, there are interdependencies between the two cycles (“connected cycles release management”, Figure 9), which we term migration/customization-integration or *MIG/CUS-integration*. The migration team must regularly test whether its scripts are still working properly with the newest customization build. For example, if the customization team decides not to offer savings books any more, the migration team cannot migrate saving books unchanged but must transform savings “books” into savings “accounts.” Workflows which the customization team adopts are another source of problems. For example, they may add a requirement that the passport-ID is stored for each new customer. If the migration team uses the same workflow for migrating existing customers, they cannot migrate customers for which they do have passport-IDs in the old system. But also the migration can cause problems for the customization team. If the migration team decides unilaterally to not migrate historic interest information, it is impossible for the customization team to generate certain tax reports.

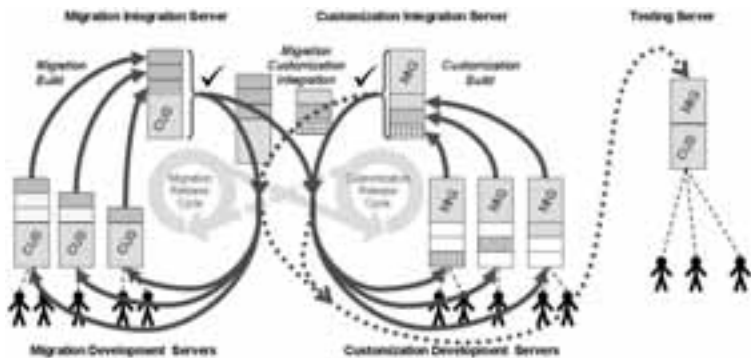


Figure 9: Connected Cycles Release Management

The dependencies require that the migration team knows the progress of the customization team and vice versa. The *connected cycles release management model* ensures exactly this. The customization team gets a “frozen” migration build from the migration team and the migration team gets a “frozen” customization build. In detail this means that the migration team works on a server with the standard software and a certain customization build (workflows and domain value tables). From time to time, the migration team runs integration tests by migrating the whole data or a subset. In case of a successful migration, the migration server has a new migration build. The migration build serves two purposes: first, it is deployed to the migration development servers as discussed before. Second, the migration test server is deployed to the customization development servers. If the workflows and domain value tables of the migration build are not the newest ones of the customization, value tables and workflows are taken from the last customization build.

Based on the new migration build, the customization team develops and adopts workflows and domain value tables. Then, they run integration tests by installing all workflow and domain value table changes on the customization integration server. This creates a new customization build. This build is deployed not only to the customization development servers. Also the migration team gets a copy as a base for their next migration build, after the already migrated data has been removed. Finally, from time to time, a copy of a migration or customization integration build is deployed to one (or more) test servers, on which testers and end users test the workflows and check the correctness of the migrated data.

7 **Quality Assurance**

Quality assurance is the third and last aspect of our data migration triangle. It subsumes three tasks:

- 1. *Migration test case identification* and their *execution* for semantic verification checks
- 2. *Reconciliation* for technical migration verification

3. *Failure reduction*, i. e. reengineering scripts which implement incorrect transformations⁵

The tasks indicate differences and common aspects between the classical testing in software development projects and the specific needs of data migration projects. This is what our risk model specific for data migration projects in Figure 10 illustrates. It is based on three assumptions. First, we assume that we do not have to deal with bugs in the old software or the new standard application. Second, we assume that the data can be migrated without prior cleansing.⁶ Finally, without loss of generality, the data of the old and the new software store their data in one database. Then, we can group the data migration specific risks as following:

- Semantic failures (Figure 10, no. 1, 2, and 3)
- Syntax errors (Figure 10, no. 4)
- Incompatibility failures (Figure 10, no. 5 and 6)

Semantic failures are conceptual problems during the extraction of data from the old system, during its transformation, or during the process of loading into the new schema. The extract step (1) must identify exactly the needed data of the old platform, which is not always easy. In case a bank migrates its customers to a new system, customers might not be of interest if they died five years ago. But if a customer caused serious problems five years ago, such that the bank decided not to do business with this customer any more, the customer must certainly be migrated and this information preserved. Secondly, the transformation might be wrong (2). Simple failures include mixing up the given name and the surname. A more complex failure would be to put all stocks of customers with more than one custody account into their first one. The third kind of semantic failures appears during the loading (3). This takes place when the migration-API of the new platform refuses certain data or data is written to wrong tables.

⁵ From a technical perspective, developing new transformations or reengineering an old one is identical. From a project management point of view it is different. The project manager, as explained in Section 4, might give more emphasis in migrating as many business objects as possible in the initial phase instead of correcting minor transformation failures.

⁶ If data cleansing is necessary, there are three options. It can be done in the old system *prior* to the migration. The advantage is that well-trained staff is available and no technical problems due to incorrect data appear during the migration. Data cleansing *during* the migration implies that the transformation solves the quality problems. It requires that all data sources needed for the cleansing are available in the system (often, they are scattered around in many Access or Excel files of different persons). However, the greatest risk is that users and managers take the chance that someone is found for the data cleansing, and then they have more and more cleansing and improvement wishes. Then, the migration team does not get along with its original work. Finally, the data cleansing can be done *after* the migration, e.g. if the project is running out of time. It requires the data quality to be in a good enough state such that it is accepted by the new system and that it does not cause serious operational problems.

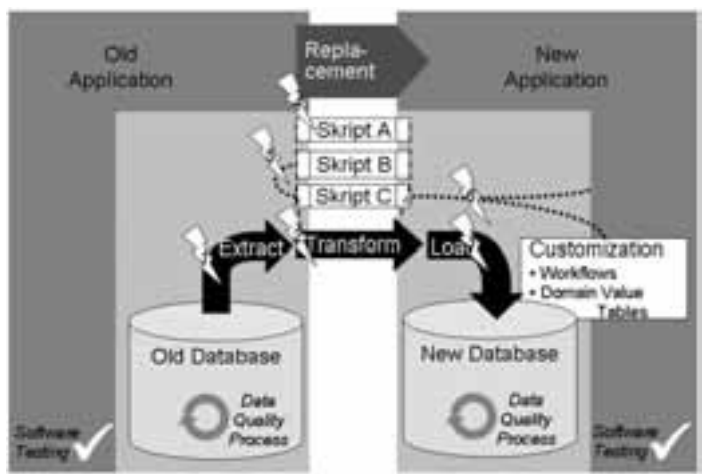


Figure 10: Risk Model for Data Migration

A failure is a *syntax error* (4) if the execution of a script raises an error due to the script not conforming to the language syntax like SQL. Finally, there are *incompatibility* failures between two migration scripts (5) or between a migration script and the standard software and its customization (6). Incompatibility failures reflect problems with MIG/MIG-integration or MIG/CUS-integration. The target platform, respectively the customization, assumes e.g. a key to be unique but the migration delivers non-unique keys. Such failures might raise run-time exceptions or the standard software refuses the data. However, if the standard software does not detect such problems immediately, they might harm stability of the software later.

We require *failure detection methods* for each of the identified risks in Figure together with KPIs as indicators for the change of the risk exposure over time (Table 1). Problems in the *extract step* are best detected using a reconciliation mechanism. *Reconciliation* means checking for each element on the source side whether it is delivered to the target, e.g. whether all customers are migrated. Furthermore, important attributes like nationality or aggregation functions like the sum of the balances of a customer’s bank accounts. User tests help in an early phase for detecting if large amounts of the data are missing (e.g. a certain branch or all savings accounts), but are of no use to check whether three out of 100,000 customers got lost. The amount of data items detected as missing (or not existing in the old system) in a certain time frame is a suitable KPI.

Transformation failures are detected best by end user tests. A bank counselor could check the portfolio of her most important customers to see whether the data in the new system has been migrated correctly. Data items identified as correct can then be checked automatically in regression tests for later migrations. A reconciliation mechanism is only of limited help, because it does not check details, and it is written by members of the migration team having less experience then users. KPIs are the number of errors during an upload (complete and/or per script) and the quantity of affected data items. Problems during the *load* are also detected by a reconciliation mechanism which might collect the data rejected by the standard software. KPI is the number of unsuccessfully migrated

data items. *Syntax errors* should be counted by the person responsible for a test migration. *Incompatibility failures* are more difficult. First, they have to be found if they do not cause a rejection during the load step. So the test migration responsible, or team members of any project team might notice such a failure. He has to report it together with a severity classification.

Data Migration Risk		Detection Method	KPI
1	Semantic failures (Extract)	+ Reconciliation o User Tests	Quantity of data set change
2	Semantic failures (Transformation)	+ User Tests o Reconciliation	Identified errors and amount of affected data elements
3	Semantic failures (Load)	+ Reconciliation	Not loaded items per test migration
4	Syntax errors	+ Report of the test migration responsible	Quantity of syntax errors overall/per script
5	Incompatibility MIG-MIG	+ Report of the test migration responsible or migration team members	Quantity and classification of severity
6	Incompatibility MIG-CUS	+ Report of the test migration responsible, migration team members, or customization team members	Quantity and classification of severity

Table 1: Failure Detection Methods and KPIs

8 Summary

When a bank replaces its core banking application platform with new (standard) software, the data of the old platform must be migrated to the new one. Such data migration projects have three dimensions, which are consolidated in our *data migration triangle*. The data migration triangle expresses and illustrates priorities for the three dimensions and their specific manifestations. In detail, the dimensions are:

- *Mapping* with the task’s business object identification, business object completeness and attribute mapping completeness. This dimension addresses that all business object types are migrated with all needed attributes.
- *Delivery* with the tasks MIG/MIG- and MIG/CUS-integration and data set completeness. The dimension represents the fact that the migration scripts not only have to run correctly for themselves, but fit together with each other and the standard software customization. Furthermore, the scripts must be executable within a sensible time frame.
- *Quality assurance* with the task of testing cases, reconciliation and failure reduction. Based on a model for technical migration risks, the dimension cares about tracking the actual risk exposure by looking how it changes over time.

With this focus on explaining the dimensions of data migration projects in companies' complex IT environments, this paper complements perfectly other work on data migration architectures like [WLB97, Hal08] by giving a management's view on the technical aspects of the actual data migration implementation.

9 References

- [Ava] Avaloq, [http:// www.avalog.com](http://www.avalog.com)
- [BA04] K. Beck, C. Andres: *Extreme Programming Explained: Embrace Change*, Addison-Wesley, Amsterdam, 2004.
- [BLW97] J. Bisbal, D. Lawless, B. Wu, et al.: *A Survey of Research into Legacy System Migration*, Technical Report TCD-CS-1997-01, Computer Science Department, Trinity College Dublin, 1997.
- [BM04] Ch. Burry, D. Mancusi: *How to plan for data migration*, in: *Computerworld*, 21.5.2004.
- [CG04] P. Carreira, H. Galhardas: *Efficient development of data migration transformations*, *Proceedings of the International Conference on Management of Data (SIGMOD)*, Paris, France, 2004.
- [DMS] *Data Management Suite (DMS)*, www.comit.ch
- [Fin] Finnova, <http://www.finnova.ch>
- [Gab07] C. Gabriel: *Plattform-Wechsel: Parforce-Übung mit weitreichenden Folgen*, in: *Schweizer Bank* (6/2007), Zürich, 2007.
- [Hal08] K. Haller: *Datenmigration bei Standardsoftware-Einführungsprojekten*, in: *Datenbank Spektrum*, Nr. 25, dpunkt.verlag, Heidelberg, 2008.
- [Hud98] J. R. Hudicka : *An Overview of Data Migration Methodology*, *Select Magazine* (April 1998), Independent Oracle Users Group, Chicago, IL, 1998.
- [Kru04] Ph. Kruchten: *The Rational Unified Process. An Introduction*, Addison-Wesley, Amsterdam, 2004.
- [LS07] *LeanStream® – COMIT Implementationsmethodik*, Version 3.0, Comit AG, Zürich, 2007.
- [Min07] M. Minetti: *Ein aktuelles Bild der Schweizer Banken IT*, www.inside-it.ch, 6.12.2007.
- [Mor06] J. Morris: *Practical Data Migration*, British Computer Society, Swindon, UK, 2006.
- [Pow] *PowerCenter*, <http://www.informatica.com/de/products/powercenter/default.htm>
- [WG04] M. Willinger, Johann Gradl: *Data Migration in SAP R/3*, Galileo Press, Boston, MA, 2004.
- [WLB97] B. Wu, D. Lawless, J. Bisbal et al., "The Butterfly Methodology : A Gateway-free Approach for Migrating Legacy Information Systems", in *Proceedings of the 3rd IEEE Conf. on Engineering of Complex Computer Systems (ICECCS'97)*, Como, Italy, 1997.

Ein Framework zur Einführung einer SOA-Governance in Unternehmen

Christian Schröpfer^{*}, Marten Schönherr^{**}

^{*}Fachgebiet Systemanalyse und EDV
Technische Universität Berlin
Franklinstr. 28/29
10587 Berlin
christian.schroepfer@sysedv.tu-berlin.de

^{**}Deutsche Telekom Laboratories
Ernst-Reuter-Platz 7
10587 Berlin
marten.schoenherr@telekom.de

Abstract: Soll die IT-Architektur eines Unternehmens auf SOA (Serviceorientierte Architektur) umgestellt werden, entstehen dadurch zahlreiche Herausforderungen auf der Managementebene. Dieser Artikel schlägt ein SOA-Management-Framework vor, welches diese Anforderungen als Änderung an der bestehenden Corporate- und IT-Governance adressiert. Es werden konkrete Gestaltungsoptionen bezüglich des Betriebsmodells, der Organisationsstruktur, der Rollen und Verantwortlichkeiten sowie der Managementprozesse dokumentiert und ein Vorgehen zur Einführung der Governance aufgezeigt. Diese Konzepte dienen als Ausgangsbasis für individuelle Einführungsprojekte in konkreten Unternehmen. Der Ansatz wurde in acht Ganztagesworkshops und einer ausführlichen Fallstudie evaluiert.

1 Einleitung

Bei der Einführung einer Serviceorientierten Architektur (SOA) in Unternehmen spielen nicht nur die technischen Aspekte der Architektur und der Serviceentwicklung eine Rolle. SOA ist vielmehr ein neues Unternehmensarchitektur-Paradigma, welches tiefgreifende Veränderungen und damit Herausforderungen auf der IT-Seite und der Fachseite mit sich bringt.

“SOA is difficult to implement, manage, and control. Not because of the technology [...] but due to the organizational, cultural, and behavioural aspects of SOA that contribute to success.” [MB06]

“When talking about enterprise IT, it is important to realize that many – if not most – of the problems associated with it are not of a technical nature but can be found on the organizational level instead.” [KBS05]

Auch eine empirische Studie der Aberdeen Group unter Unternehmen, die eine SOA einführen bzw. dies planen, bestätigt diese Aussagen [KS07].

Einer der Hauptgründe liegt in der hochverteilten Architektur, welche auf vielen verschiedenartigen, mehrfach verwendeten und flexiblen Services aufbaut und durch komplizierte Nutzer-Anbieter-Verhältnisse über organisatorische Grenzen hinweg gekennzeichnet ist. Die Services befinden sich auf unterschiedlichen Granularitätsstufen und in unterschiedlichen Lebenszyklusphasen. Diese Situation birgt zahlreiche Risiken. Zum einen kann der Überblick über das Gesamtsystem leicht verloren gehen. Zum anderen kann der Ausfall auch nur eines Services negative Auswirkungen auf verschiedene wichtige Geschäftsprozesse haben.

Den Herausforderungen muss sich ein Unternehmen durch zahlreiche größere und kleinere Veränderungen bezüglich Organisation und Managementabläufen insbesondere im IT-Bereich – aber auch Fachbereich – stellen. Nur so können die Vorteile, die potenziell mit einer SOA verbunden sind, genutzt werden. Unter anderem ist es notwendig, ein zentrales und konsistentes Servicemanagement aufzubauen, die Wiederverwendung zu forcieren, die stabile Bereitstellung von Services über organisatorische Grenzen hinweg sicherzustellen, die Risiken zu managen sowie die Services strikt an den Bedürfnissen der Geschäftsseite auszurichten.

Die Herausforderung besteht insbesondere darin, den SOA-spezifischen Anforderungen mit möglichst wenig Veränderung gegenüber generischen (hoffentlich im Unternehmen bereits etablierten) IT-Governance-Ansätzen gerecht zu werden. Das SOA-Management-Framework bündelt diese notwendigen Veränderungen und Maßnahmen auf organisatorischer Ebene zu einem praktikablen und einsatzbereiten Konzept und unterstützt damit Unternehmen bei der erfolgreichen Einführung und dem nachhaltigen Betrieb einer SOA entscheidend. Es wurde bereits mit verschiedensten Firmen ausführlich diskutiert und dabei verfeinert.

Der Artikel ist wie folgt gegliedert: Zunächst werden die theoretischen Grundlagen der SOA, der IT-Governance und der SOA-Governance diskutiert. Danach wird ein Überblick über das SOA-Governance-Framework mit Details zum Vorgehensmodell bei der Einführung und in den Bereichen Betriebsmodell und Organisationsstruktur sowie Managementprozesse und Richtlinien gegeben. Danach wird die Methodik beschrieben, mit der das Framework erarbeitet und in der Praxis validiert wurde.

2 Theoretische Grundlagen

In diesem Abschnitt werden die Grundlagen der SOA, der IT-Governance sowie vorhandene IT-Governance-Frameworks und SOA-Governance-Ansätze beschrieben.

2.1 SOA

Für den Begriff SOA existiert keine einheitliche Definition. Die Bandbreite reicht von nur technisch geprägten Definitionen bis hin zu Definitionen, die das gesamte Unternehmensmanagement mit einbeziehen. Der Begriff SOA lässt sich nach einer eigenen Definition folgendermaßen definieren:

Die SOA ist ein Architekturparadigma, welches Elemente der Softwarearchitektur und der Unternehmensarchitektur miteinander vereint. Es basiert auf der Interaktion von Services, die autonom und interoperabel sind und fachlich relevante, wiederverwendbare Funktionen über eine technisch standardisierte Schnittstelle anbieten. Services können auf allen Anwendungssystemsichten (Präsentation, Geschäftsprozess, Geschäftslogik, Datenhaltung) existieren, aus Services tieferer Schichten zusammengesetzt und aus bestehenden Anwendungssystemen gekapselt, aber auch neu implementiert sein.

Zu den allgemein anerkannten Charakteristiken serviceorientierter Architekturen zählen weiterhin die verteilte Struktur sowie Aspekte der Orchestrierung und losen Kopplung von Services. Teilaspekte dieser Definition finden sich in der von McCoy und Natis sowie der von Marks und Bell wieder [MN03; MB06].

In diesem Artikel wird von folgendem Servicebegriff ausgegangen: „Ein Service stellt ein abstraktes Software-Element bzw. eine Schnittstelle dar, die anderen Applikationen über ein Netzwerk einen standardisierten Zugriff auf Anwendungsfunktionen anbietet.“ [He07] Der verwendete Servicebegriff ist damit von den klassischen IT-Services (wie in ITIL verwendet, z.B. Desktopservice, Netzwerkservice) klar abgegrenzt.

Die Einführung einer SOA im Unternehmen wird mit zahlreichen Vorteilen verbunden: Wiederverwendung, damit verbunden Kosteneffizienz und gesteigerte Qualität der Software(-komponenten), Flexibilität, d.h. schnellere und einfachere Anpassung der IT-Systeme an die (geänderten) Geschäftsprozesse, gesteigerte Transparenz und Outsourcing-Möglichkeiten, niedrige Einführungszeit neuer Produkte bzw. Services (engl. Time-to-Market) und hohe Kompatibilität. Als weitere Ziele werden ein effizienterer Entwicklungsprozess, adäquate Geschäftsinfrastruktur, Möglichkeit der schrittweisen Einführung und Erweiterung, Risikominderung und Unabhängigkeit von der Technologie genannt. All diese sind Charakteristiken der agilen Unternehmung [Wo03; MB06]. Eine geeignete SOA-Governance ist eine Voraussetzung für die Realisierung der genannten Ziele.

2.2 IT-Governance

Zunächst wird ein Verständnis des Begriffes „SOA-Governance“ basierend auf Definitionen der Begriffe „Corporate-Governance“ und „IT-Governance“ erarbeitet.

„... corporate governance could be defined as ‘the process of controlling management and of balancing the interests of all internal stakeholders and other parties (external stakeholders, governments and local communities [...]) who can be affected by the corporation’s conduct in order to ensure responsible behavior by corporations and to achieve the maximum level of efficiency and profitability for a corporation’.” [PMB05]

Aufgrund der Bedeutung der IT für heutige Unternehmen kann Corporate-Governance nicht unabhängig vom Management der IT betrachtet werden. Vielmehr muss die IT-Governance in der Corporate-Governance integriert werden [Gr03]. IT-Governance ist der Teil der Governance, der sich auf die Informations- und Kommunikationssysteme im Unternehmen bezieht. Das ITGI (IT Governance Institute), Autor des COBIT-Frameworks (Control Objectives for Information and related Technology), definiert IT-Governance wie folgt:

„IT governance is the responsibility of executives and the board of directors, and consists of the leadership, organisational structures and processes that ensure that the enterprise’s IT sustains and extends the organisation’s strategies and objectives.“ [IT07]

Peterson definiert IT-Governance folgendermaßen:

„The distribution of IT decision-making rights and responsibilities among enterprise stakeholders, and the procedures and mechanisms for making and monitoring strategic decisions regarding IT.“ [Pe04]

Es geht nicht darum, wie einzelne Entscheidungen gefällt werden, sondern um die Ausgestaltung der Entscheidungsrechte und der Zuständigkeiten [RW04]. Das Verhalten in der Benutzung der IT soll zielgerichtet beeinflusst werden. Davon ausgehend lässt sich der Begriff SOA-Governance definieren:

„SOA governance is more than providing governance for SOA efforts; it is how IT governance should operate within an organization that has adopted SOA as their primary approach to enterprise architecture.“ [BI04]

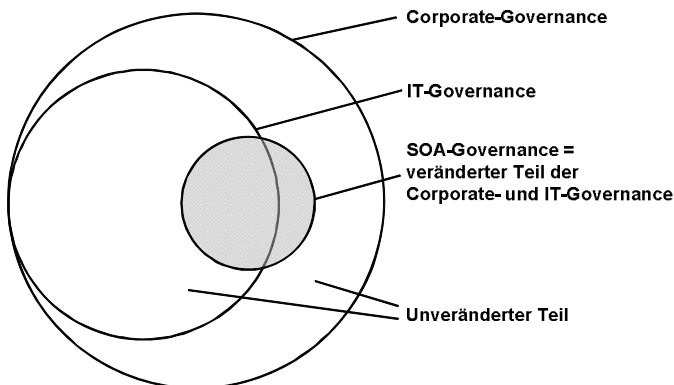


Abbildung 1: Verhältnis von Corporate-Governance, IT-Governance und SOA-Governance

2.3 IT- und SOA-Governance-Frameworks

Das hier eingeführte SOA-Governance-Framework basiert auf zwei Fundierungen: IT-Governance-Literatur und SOA-spezifischer Managementliteratur.

In einem Vergleich aktueller IT-Governance-Literatur in der Zeitschrift „Wirtschaftsinformatik“ [Bu05] zwischen fünf Titeln werden die Werke [Gr03], [RW04] und [IT03] inhaltlich besonders gut bewertet. Die Grundlage des hier vorgestellten SOA-Governance-Frameworks basiert auf den ersten beiden Quellen und dem CobiT-Framework [IT07], welches in einer älteren Version der letzten Quelle zugrunde liegt.

Nach van Grembergens ganzheitlichem Framework gehören zur IT-Governance die folgenden Blöcke: Strukturen, Prozesse und Beziehungsmechanismen.

„Structures involve the existence of responsible functions such as IT executives and accounts, and a diversity of IT committees. Processes refer to strategic IT decision-making and monitoring. The relational mechanisms include business/IT participation and partnerships, strategic dialogue and shared learning.” [Gr03]

Ross und Weill präsentieren ein „Governance Design Framework“ der MIT Sloan School for IS Research mit den zu harmonisierenden Komponenten [RW04]:

- Unternehmensstrategie und Organisation
- IT-Organisation und wünschenswertes Verhalten
- IT-Governance-Gestaltungskomponenten
- IT-Governance-Mechanismen mit IT-Entscheidungen bezüglich Prinzipien, Architektur, Infrastruktur und Investitionen
- Geschäftsbezogene Leistungsziele
- IT-Metriken und Verantwortlichkeiten

Die von van Grembergen sowie Ross und Weill genannten Aspekte finden sich im erarbeiteten Framework in den Bereichen Betriebsmodell und Organisationsstruktur (mit Rollen und Verantwortlichkeiten), Managementprozesse und Richtlinien, Zusammenarbeit und Kommunikation sowie Performance Management wieder (vgl. Abb. 1).

CobiT ist ein praxisorientiertes Framework zur Implementierung von IT-Governance [IT07]. Es dokumentiert 34 Prozesse in den vier Prozessbereichen „Plan and Organise“, „Acquire and Support“, „Deliver and Support“ und „Monitor and Evaluate“. Zu jedem Prozess werden die Teilaktivitäten mit Beschreibungen, Inputs, Outputs, RACI-Tabellen („Responsible“, „Accountable“, „Consult“, „Inform“) und Ziel- und Performanceindikatoren angegeben. Diese Prozesse sollen ein effektives Management der Informationen und IT-Ressourcen gewährleisten [SS07b].

Auch ITIL (IT Infrastructure Library) des ITSMF (IT Service Management Forum) beschreibt Best Practices für IT-Managementprozesse und ist ein etablierter Standard. ITIL ist mittlerweile in der Version 3 erschienen [OGC07]. Die Dokumentation besteht aus den fünf Kernbereichen „Service Strategy“, „Service Design“, „Service Transition“, „Service Operation“ und „Continual Service Improvement“, die nach dem IT-Servicelebenszyklus gegliederte Gestaltungsvorschläge bezüglich der Organisation von Service Providern enthalten.

Das hier vorgestellte SOA-Governance-Framework greift die Prozesse von ITIL und CobiT heraus, die SOA-spezifische Änderungen erfahren müssen. Diese gezielten Änderungen sind leichter zu implementieren, wenn im Unternehmen bereits eine IT-Governance (CobiT) und adäquates IT-Service-Management (ITIL) realisiert sind. Die SOA-spezifischen Managementprozesse sind mit den CobiT-relevanten Informationen angereichert.

In der Literatur wird das Problem SOA-Governance bereits diskutiert. Der Artikel [Bi05] beschreibt einen „Human Service Bus“ als richtige organisatorische Aufstellung in einer Organisation, die SOA einführt. Er schlägt eine domänenbasierte Organisationsstruktur vor. In Domänen werden Dienste mit ähnlichen fachlichen Anforderungen zusammengefasst und servicebezogene Entscheidungen und Managementaktivitäten ausgeführt. Der Artikel gibt zudem allgemeine Hinweise, wie Mitarbeiter durch gezieltes auf individuelle Anreize ausgerichtetes Performance Management positiv beeinflusst werden können. Obwohl der Artikel für Teile des hier vorgestellten Frameworks eine gute theoretische Basis darstellt, gibt er keine konkreten Gestaltungsvorgaben an. In [Bi06], u.a. vom selben Autor verfasst, werden einzelne Rollen (ohne KPIs) dokumentiert.

[Sy06] ist ein Whitepaper und hebt die Bedeutung der SOA-Richtlinien, der Serviceverträge, des Metadatenmanagements und des Servicelebenszyklusmanagements hervor. Insbesondere wird der zweigeteilte Provider- und Konsumentenlebenszyklus dargestellt. Der Artikel gibt über die Motivation zur Beschäftigung mit der SOA-Governance und die Darstellung des Servicelebenszyklus hinaus keine weiteren Designvorgaben, die bei der Umsetzung einer SOA-Governance helfen würden. [SS06a] stellt ebenfalls die Notwendigkeit für ein SOA-Governance-Framework, ein SOA-Metamodell und kurz eine beispielhafte Entwicklung eines SOA-Governance-Modells vor.

[ST07] enthält verschiedene Artikel im Bereich Geschäftsaspekte der SOA und SOA-Governance. Die Einzelartikel sind jedoch für sich abgeschlossen und bauen nicht aufeinander auf. [Wo06] betrachtet folgende Teilbereiche der SOA-Governance: Servicedefinition, Serviceeinbetriebnahme, Servicelebenszyklus, Serviceversionierung, Servicemigration, Serviceregistries, Servicedatenmodelle, Serviceüberwachung, Serviceeigentum, Servicetests und Servicesicherheit. Teilweise werden Lösungsansätze beschrieben, die im vorgestellten Framework verarbeitet sind. [BI04] beschreibt auf grober Ebene Architektur-, Geschäfts-, Daten- und technische Prinzipien, die in einer SOA verwendet werden sollen, dokumentiert, welche Architekturmanagementprozesse notwendig sind und hebt das Domänenprinzip hervor.

Im Bereich Architekturmanagement in einer SOA wird in der Open Group (www.opengroup.org/projects/soa-togaf) daran gearbeitet, wie TOGAF 8 (The Open Group Architecture Framework) beim Management einer SOA konkret eingesetzt werden kann. Erste Ergebnisse werden in [DH08] diskutiert.

Die dargestellte IT- und SOA-Governance-Literatur bildet die Grundlage für das vorgestellte SOA-Governance-Framework. Aus diesen Quellen wird im Folgenden die Struktur des SOA-Governance-Frameworks abgeleitet und dort dargestellte Inhalte in die Struktur integriert.

3 SOA-Governance-Framework

Für die SOA-Governance relevante Themengebiete sind das Unternehmensarchitekturmanagement, das Geschäftsprozessmanagement, das Servicemanagement, das Management von Programmen und Projekten, das Providermanagement und das Infrastrukturmanagement (vgl. Abb. 2). Diese Themenbereiche lassen sich durch in [RWR06] durchgeführte empirische Untersuchungen belegen. [RWR06] analysiert Herausforderungen und Lösungsmöglichkeiten in der Unternehmensarchitektur und angrenzenden Bereichen basierend auf empirischen Forschungsprojekten mit etwa 200 Unternehmen, die den vier Reifegradstufen „Unternehmenssilos“, „Standardisierte Technologie“, „Optimierter Kern“ und „Modulares Unternehmen“ zugeordnet wurden. Auf dem Weg zum modularen Unternehmen, für welches SOA eine Realisierungsvariante ist, werden verschiedene Herausforderungen genannt, die sich den oben als relevant genannten Bereichen zuordnen lassen: Unternehmensarchitekturmanagement, unternehmensweite IT-Governance, Identifikation und Realisierung der Bedürfnisse von der Fachseite und der Geschäftsprozesse, Management von wiederverwendbaren, standardisierten Geschäftskomponenten, professionelle Projektmanagementfähigkeiten, Unternehmensarchitektur als Richtlinie für Outsourcing-Modelle, Management der IT-Infrastruktur.

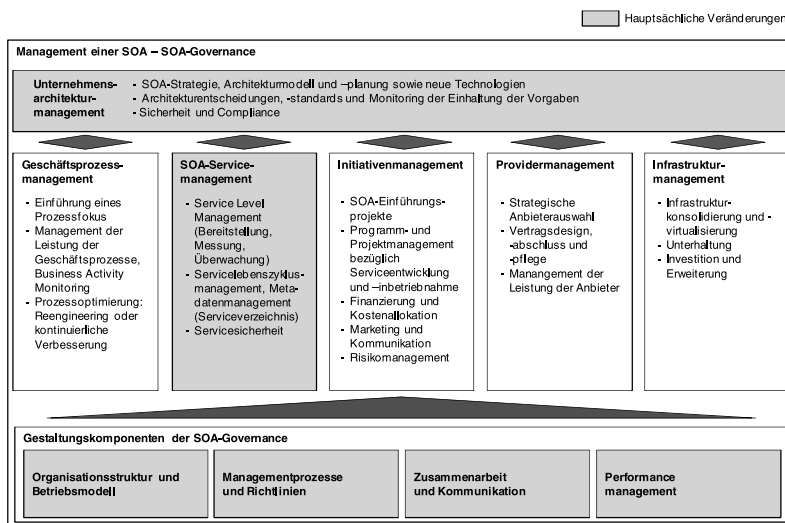


Abbildung 2: Überblick über die Aspekte der SOA-Governance

Das SOA-Management-Framework konzentriert sich auf die beiden Kernbereiche Architektur- und Servicemanagement. Im Rahmen des Architekturmanagements sind unter anderem eine SOA-Strategie, Architekturarten und eine Architekturroadmap zu erarbeiten sowie Architekturstandards festzulegen und nachzuhalten. Bestandteile des Servicemanagements sind das Management der Services entlang ihres Lebenszyklus und das Management der Servicelevels (Dienstgütevereinbarungen). Die diesbezüglichen Vorschläge zum Design eines SOA-Governance-Konzeptes werden von einem Vorgehensvorschlag zur Einführung einer SOA-Governance begleitet, der den jeweiligen Reifegrad des Unternehmens berücksichtigt.

3.1 Vorgehen bei der Einführung einer SOA-Governance

Abb. 3 zeigt das vorgeschlagene Vorgehen bei der Einführung einer SOA-Governance in einem Unternehmen. Zunächst werden eine Strategie und eine grobe Roadmap bezüglich der SOA-Einführung erarbeitet, die SOA-Ziele definiert, das Team zur Erarbeitung der SOA-Governance zusammengestellt und das generelle Vorgehen abgestimmt. In der Analysephase werden der Reifegrad bezüglich der IT-Governance und der SOA sowie die allgemeine Ausgangssituation und Veränderungsbereitschaft des Unternehmens untersucht. Die Analyse wird mit Hilfe eines Fragebogens durchgeführt, der aus drei Teilen besteht:

- Reifegrad der IT-Governance nach dem Reifegradmodell von CobiT [IT07]
- Reifegrad der IT-Managementprozesse nach CMMI [Cm06]
- SOA-spezifischer Reifegrad nach individuell entwickelten Fragen zum Stand der SOA und bereits durchgeführten Projekten sowie etablierten SOA-Governance-Strukturen

In der Designphase werden die Governance-Vorgaben und -Mechanismen unter Berücksichtigung dieser Analyse an die spezifische Situation im Unternehmen angepasst und dokumentiert. Das im nächsten Abschnitt vorgestellte SOA-Governance-Konzept gibt dazu Designvorlagen und -vorschläge. Die Schritte zur Implementierung dieser Konzepte werden in der Planungsphase auf eine Zeitachse gelegt und Meilensteine definiert. Trainingsmaterial wird entwickelt. Die Meilensteine werden in der darauf folgenden Implementierungsphase zur Fortschrittskontrolle verwendet. Die Mitarbeiter mit veränderten Aufgabenbereichen werden geschult. In der letzten Phase, der Optimierungsphase, werden die Managementabläufe und die Güte der SOA überwacht, bei Abweichungen vom Soll werden Veränderungen durchgeführt. Die Governance-Konzepte und Anzahl der Mitarbeiter pro Rolle werden angepasst.



Abbildung 3: Vorgehen bei der Einführung einer SOA-Governance

3.2 Überblick über die Basiskonzepte für die Designphase

Alle beschriebenen Konzepte sind generische Vorschläge für Änderungen an der derzeitigen IT-Governance bei der Einführung einer SOA-Governance. Sie wurden im Rahmen des theoretischen Literaturstudiums erarbeitet und in Experteninterviews, in Workshops und in einer Fallstudie validiert und verbessert. Die Konzepte stellen im Einzelfall keine mandatorische Zielstruktur, sondern eine Ausgangsbasis für die Diskussion zur unternehmensspezifischen Ausgestaltung der SOA-Governance dar.

Von den vier oben abgeleiteten Bestandteilen des Frameworks zur Umsetzung von Architektur- und SOA-Servicemanagement (vgl. Abb. 4) werden hier Betriebsmodell und Organisationsstruktur sowie die Managementprozesse und Richtlinien näher behandelt.

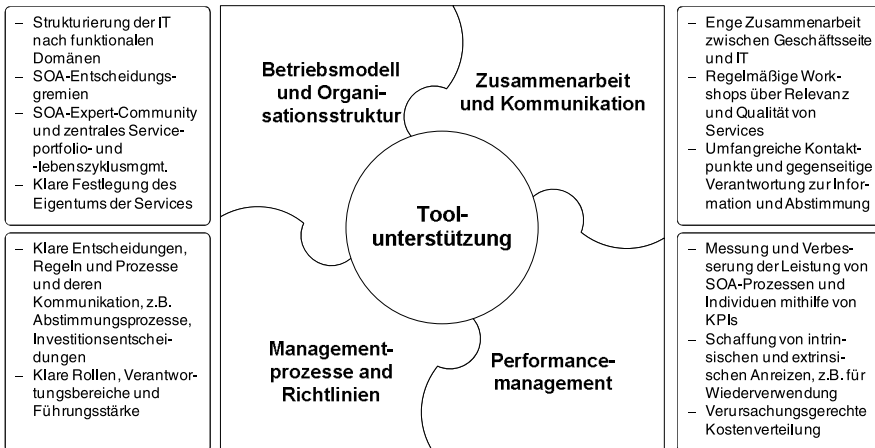


Abbildung 4: Gestaltungscomponenten der SOA-Governance – in Anlehnung an [Gr03]

3.3 Organisationsstruktur und Betriebsmodell

In diesem Bereich sind potenziell vier Arten von Anpassungen notwendig, die Anpassung des Betriebsmodells mit der Organisationsstruktur auf Geschäfts- und IT-Seite sowie die Einführung von virtuellen SOA-Strukturen, von SOA-spezifischen Verantwortlichkeiten zu existierenden Rollen und von neuen SOA-spezifischen Rollen – soweit noch nicht existent.

Die notwendige Anpassung des Betriebsmodells und damit der Organisationsstruktur hängt maßgeblich von der initialen Aufstellung des Unternehmens ab. Während der Erarbeitung und Validierung des Frameworks wurden basierend auf theoretischem Literaturstudium und aus den angetroffenen realen Situationen in den untersuchten Unternehmen fünf verschiedene Ausgangssituationen abgeleitet. Diese sind:

1. großer IT-Service-Provider ohne IT-Demand-Organisation
2. großer IT-Service-Provider mit einer IT-Demand-Organisation
3. sehr kleine IT-Einheit ohne vorherige SOA-Erfahrung
4. großer IT-Service-Provider ohne eine IT-Demand-Organisation, welche international verteilte Geschäftsstandorte mit ähnlichem Geschäft bedient
5. großer IT-Service-Provider ohne eine IT-Demand-Organisation, welche national verteilte Geschäftsstandorte mit demselben Geschäft bedient

Als IT-Demand-Organisation wird in diesem Artikel eine flache IT-Abteilung verstanden, die die IT-Anforderungen der Fachseite bündelt, in technische Anforderungen umsetzt und die IT-Abteilungen (intern oder extern), die die eigentlichen IT-Services erbringen, steuern. Für Ausgangssituation 1 wird nun beispielhaft ein SOA-Betriebsmodell beschrieben (vgl. Abb. 5).

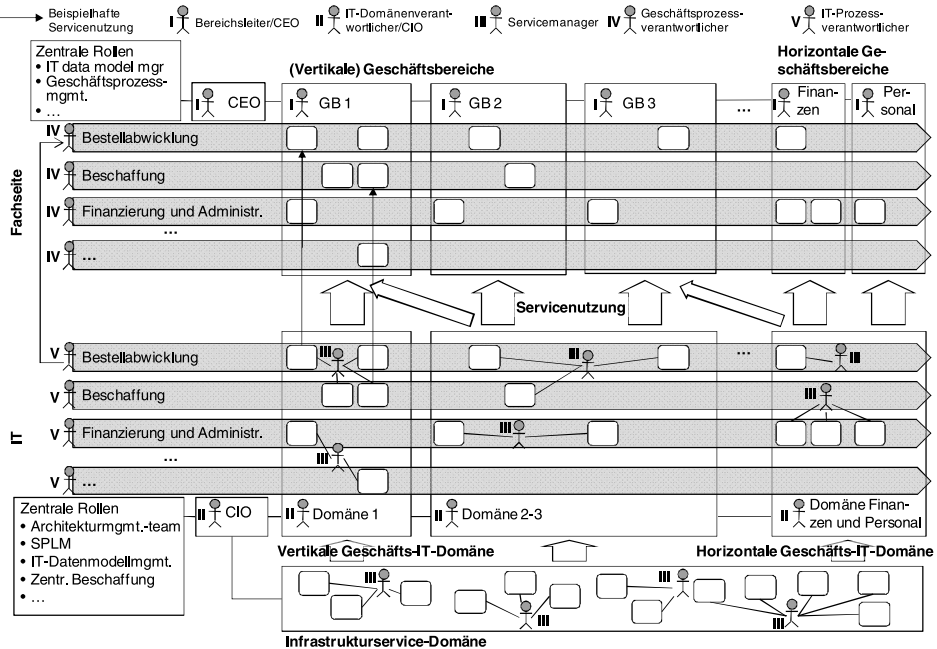


Abbildung 5: Überblick über das SOA-Betriebsmodell für Ausgangssituation 1

Bei den Änderungen am Betriebsmodell wurden folgende Designziele berücksichtigt:

1. Funktionale Expertise auf der Fachseite und der IT-Seite sowie Wiederverwendung von Services und anderen Artefakten
2. Prozessfokus auf der Fachseite und der IT-Seite
3. Serviceorientierung
4. Geschäftsseite als Treiber für IT-Funktionalität
5. Technisches Wissen auf der IT-Seite
6. Unterstützung der jeweiligen Geschäftsstrategie und Flexibilität hinsichtlich des Geschäftsmodells einzelner Abteilungen

Begründet durch die Designziele 1. und 2. muss das Betriebsmodell im Ziel eine Art Matrixorganisation bilden, die funktional strukturierte Einheiten enthält und gleichzeitig einen Prozessfokus hat.

Die funktionale Expertise macht eine funktionale Strukturierung der Fachseiten notwendig. Diese Strukturierung der Organisation, die auf der funktionalen Spezialisierung und damit dem Aufbau von funktionalen Kompetenzen beruht, ist deswegen heute schon weit verbreitet, z.B. nach Produkten, Servicelinien und Kundengruppen. Den funktionalen Einheiten auf der Fachseite stehen Bereichsleiter vor, die volle Geschäftsverantwortung für ihren Bereich innehaben. Die Bereichsleiter führen zusammen mit den Abteilungs- und Unterabteilungsleitern die Mitarbeiter, die die IT-Systeme entlang der

Geschäftsprozesse zur Bearbeitung ihrer geschäftlichen Aufgaben einsetzen. Aus der Geschäftsarchitektur können funktionale Domänen abgeleitet werden, die nicht unbedingt mit den organisatorischen Einheiten auf der Fachseite übereinstimmen müssen.

Auch auf der IT-Seite soll funktional spezifisches Wissen, welches zur Entwicklung der benötigten IT-Funktionalität z.B. in Services notwendig ist, nicht jedesmal wieder von Neuem aufgebaut werden müssen. Darüber hinaus sollen die Services in der Organisation wiederverwendet werden (dicke Pfeile in Abb. 5). Deshalb sollte die Organisation auf der IT-Seite nach den funktionalen Domänen strukturiert sein und sogenannte IT-Domänen und -Unterdomänen gebildet werden. Es gibt horizontale und vertikale IT-Domänen sowie eine Infrastrukturservicedomäne. In vertikalen Domänen werden Services mit spezifischer Fachfunktionalität gebündelt, die für einen Teil der fachlichen Organisationseinheiten nützlich sind. In horizontalen Domänen werden Services für fachliche Querschnittsfunktionen gebündelt. In der Infrastrukturservicedomäne werden Infrastrukturservices angeboten, also beispielsweise Services zur Verwaltung von Ressourcen oder zur Steuerung der Sicherheit oder zur Überwachung.

Die IT-Abteilungen werden von IT-Domänenverantwortlichen geführt, die für die gesamte Erbringung der Dienstleistungen ihrer Abteilung und damit insbesondere für die Erbringung der SOA-Services verantwortlich sind und an den Leiter der IT (CIO) berichten. Die Servicemanager sind für kleinere Gruppen von SOA-Services innerhalb der Domänen verantwortlich.

Der Prozessfokus spielt in einer SOA eine besondere Rolle. Er wird sowohl auf der Fachseite als auch der IT-Seite durch Prozessverantwortliche realisiert. Der Prozessverantwortliche auf der Fachseite ist für den Ablauf und die Verbesserung der Geschäftsprozesse (Geschäftsprozessmanagement) verantwortlich. Dadurch soll die Flexibilität der SOA ausgenutzt werden und echter Mehrwert für die Fachseite geschaffen werden. Der IT-Prozessverantwortliche hat technisches Wissen über die aufgerufenen Services, deren Schnittstellen und Orchestrierung. Er unterstützt den fachlich Verantwortlichen bei der technischen Realisierung eines veränderten Prozesses und bei der Fehlersuche. Beide können mehrere Prozesse verantworten. Dem Autorenteam ist klar, dass es eine besondere Herausforderung darstellt, einen Prozessfokus in einem Unternehmen einzuführen. Probleme können u.a. eine schwache Stellung der Prozessverantwortlichen gegenüber den funktional orientierten Stellen und der Verlust der Zugehörigkeit zu einer Hierarchie bei den Prozessverantwortlichen sein. Die Prozessverantwortlichen sollen dem CEO im Rahmen einer zentralen Geschäftsprozessmanagementeinheit berichten, die IT-Prozessverantwortlichen dem CIO als Teil einer IT-Prozessabteilung.

Gemäß den Designzielen „Serviceorientierung“ und „Geschäftsseite als Treiber für IT-Funktionalität“ geht der Bedarf für neue Funktionalität von der Fachseite aus und die IT-Abteilung des Unternehmens funktioniert als Serviceprovider für die Fachseite. Letzten Endes ist die Fachseite für neue Funktionalität verantwortlich und beauftragt diese.

Ein Kernpunkt bei der Einführung eines neuen Paradigmas und einer neuen Technologie ist der Aufbau und die Migration der Kompetenzen in der Organisation. Die zusätzlichen Ressourcen in einem Unternehmen sind häufig begrenzt. Die SOA-Expert-Community

ist eine virtuelle Einheit, der interessierte Mitarbeiter aus dem Architekturteam und den IT-Domänen angehören. In dieser werden service-, architektur- und technologiebezogenes Wissen sowie Erfahrungen ausgetauscht. Die Mitglieder bekommen einen gewissen – insbesondere zeitlichen – Freiraum, in dem dieser Wissensaustausch und die Wissens-erarbeitung stattfinden können. Gefördert werden der Zusammenhalt und der Wissensaustausch durch persönliche Gruppentreffen. Die Wissensweitergabe ist fester Bestandteil des Beförderungsmechanismus und führt zu positiver Sichtbarkeit im Unternehmen. Die SOA-Experten beeinflussen die Gesamtarchitektur einer Lösung am Anfang oder an schwierigen Stellen. Sie tragen ihr Wissen von einem realen Projekt zum nächsten und kontaktieren sich untereinander, wenn ihnen die spezifische Expertise fehlt.

Zahlreiche Rollen müssen neu eingeführt oder von ihrer Verantwortung her angepasst werden. In Tabelle 1 sind die wichtigsten dargestellt. Zu den neuen Rollen existieren Rollenbeschreibungen mit Ziel, Vorgesetztem, KPIs, benutzten Tools, besonderen Fähigkeiten, Mitarbeit in Gremien, Anzahl der Mitarbeiter und RACI-Angaben.

Weitere Gremien und Rollen, die bereits existieren und einen leicht veränderten Aufgabenbereich haben, sind der Geschäftsbereichsleiter, das IT-Strategie-Komitee, der IT-Lenkungsausschuss, das Providermanagement, der Projektmanager, Business-Analysten, Softwareentwickler, Sicherheitsspezialisten, Deployment-Manager, Testspezialisten, Applikationsverantwortliche sowie System- und Datenbankadministratoren.

Rolle/Gremium	(N)eu/ (V)er- ändert	Hauptsächliche Aufgaben	KPIs
IT-Domänen-verantwortlicher	N	Wirtschaftliche und fachliche Verwaltung der SOA-Services der horizontalen oder vertikalen Fachdomäne bzw. der Infrastrukturservicedomäne	Wiederverwendungsrate, SLA-Konformität der Services, Zufriedenheit der Service-nutzer
Servicemanager	N	Management einer Gruppe von Services in einer Domäne entlang des gesamten Lebenszyklus; Ansprechpartner für Fachseite bezgl. IT-Funktionalität; Verantwortung für Verhandlung und Einhaltung der SLAs	Wiederverwendungsrate, SLA-Konformität der Services, Zufriedenheit der Service-nutzer
Geschäftsprozess-verantwortlicher	N	Verbesserung der Performanz der Geschäftsprozesse	Prozessspezifische KPIs, wie Durchlaufzeit
IT-Prozessverantwortlicher	N	Vermittlung zwischen Geschäftsprozessverantwortlichen und Servicemanagern; Lösung von technischen Problemen mit dem Prozess	Prozessspezifische KPIs auf technischer Ebene, z.B. Fehlerraten
Architektur Review Board	N	Kontrolle der Projekte hinsichtlich der Einhaltung der Architektur- und Technologierichtlinien; Freigabe von Services	-

Zentrales Architekturteam	V	Entwicklung von IT-Architektur- und Technologierichtlinien; Entsendung von Architekten in Projekte	-
Geschäftsprozessmanagement	N	Methodische Unterstützung bei der Aufnahme von Prozessen; organisatorische Heimat der Geschäftsprozessverantwortlichen	-
Zentrales Serviceportfolio- und -lebenszyklusmanagement	N	Koordination von Serviceanfragen; Management des SOA-Serviceportfolios; Qualitätskontrolle der Servicemetadaten und Unterstützung der Wiederverwendung	-
Datenmodellmanager	N	Initiale Entwicklung eines unternehmens-/domänenweiten logischen Datenmodells und Erweiterung	Konsistenz und Abdeckungsgrad des logischen Datenmodells
Technischer Prozessdesigner	N	Erweiterung der modellierten Geschäftsprozesse durch Fehlerbehandlung, Sicherheitsaspekte und Transaktionalität	Prozessverfügbarkeit, Prozessfehlerquote
Softwarearchitekt	V	Definition der Architektur des zu entwickelnden IT-Systems nach den Richtlinien; Definition der Datentypen, Schnittstellen, Nachrichtenschemata und -austauschmuster der Services	Qualität der Architektur, Wiederverwendung von Services, Servicequalität

Tabelle 1: Tabelle mit neuen bzw. veränderten Rollen sowie deren Aufgaben und KPIs

3.4 Managementprozesse und Richtlinien

Das SOA-Management-Framework sieht für diese Rollen zahlreiche Managementprozesse und Richtlinien vor. Richtlinien, die die Kommunikation, den Zugriff auf die Services, Kontrollmechanismen und das Verhalten der Mitarbeiter regeln, sind für alle SOA-spezifischen Bereiche des Architektur- und Servicemanagements enthalten. Im Einzelnen sind dies:

Architekturmanagement

- Funktionales Domänenmodell entwickeln
- Anwendungslandschaft dokumentieren
- Architekturroadmap planen
- Strategisches Service-/Projektportfolio priorisieren
- Initiales logisches Datenmodell entwickeln
- Logisches Datenmodell auf Anfrage erweitern

SOA-Service-Management

- Servicedesign überprüfen
- Serviceversionierung und -änderungsmanagement durchführen

- Service-Levels planen und verhandeln
- Service Levels überwachen
- Services von externem Provider nutzen

Zusätzlich zu den Managementprozessen müssen in das eigentliche Vorgehensmodell zum Design der Services bestimmte Managementaktivitäten eingebaut werden, z.B. Test und Suche nach wiederverwendbaren Services, Codebestandteilen, Anforderungen und Architekturplänen, Architekturreviews des Servicedesigns, Ermittlung der Anforderungen anderer potenzieller Servicenutzer, deren Priorisierung und Aggregation, ausführliche Tests (Serviceeinzeltests, Prozess- und Serviceintegrationstests) sowie Abschließen eines formalen Serviceerstellungs- und Betriebsvertrages.

Beispielhaft wird hier der SOA-Serviceversionierungs- und -änderungsmanagement-Prozess dargestellt (siehe Abb. 6). Dieser ist ein Kernbestandteil des Servicelebenszyklusmanagements. Servicelebenszyklusmanagement ist die Koordination und Planung der mit den Diensten zusammenhängenden Aktivitäten in allen Phasen des Lebenszyklus: Identifikation und Planung, Erarbeitung des Konzeptes, Anforderungsdefinition, Servicedesign, Implementierung, Einführung und Inbetriebnahme, Betrieb, Außerbetriebnahme sowie insbesondere Updates und Migrationen der Services.

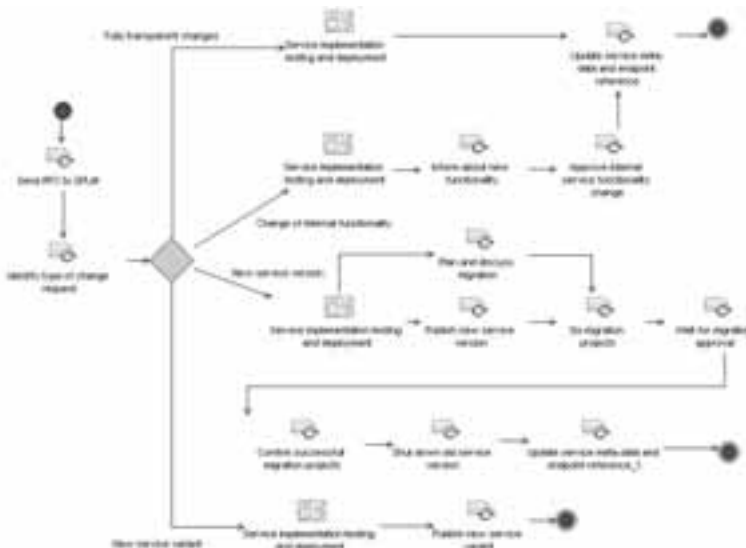


Abbildung 6: Prozess SOA-Serviceversionierungs- und -änderungsmanagement¹

¹ Dargestellt in der proprietären Prozessnotation des Eclipse Process Frameworks (<http://www.eclipse.org/epf/>)

Dieser Prozess beschreibt, wie existierende Services mit Hilfe von Versions- und Variantenbildung geändert werden können, um die Stabilität des gesamten IT-Systems zu erhalten, d.h. bestehende Servicenutzungen nicht zu gefährden. Der Prozess wird nach dem Servicedesign gestartet, wenn existierende relevante Services geändert werden sollen. KPIs, mit denen die Leistung dieses Prozesses überprüft werden kann, sind u.a. Prozentsatz der Änderungen, die durch diesen Prozess erfasst werden, Schnelligkeit der Entscheidung sowie die Anzahl der pünktlich und erfolgreich abgeschlossenen Migrationsprojekte.

Vier Hauptfälle müssen unterschieden werden

- Für Nutzer voll transparente (d.h. keine relevanten funktionalen) Änderungen. In diesem Fall muss der geänderte Service lediglich entwickelt, getestet, in Betrieb genommen und im Repository aktualisiert werden.
- Änderung interner Funktionalität, jedoch ohne Änderungen am Interface: In diesem Fall müssen die Nutzer über die neue Funktionalität informiert werden und dieser zustimmen.
- Neue Serviceversion: Änderungen der Funktionalität und Schnittstelle. Die neue Funktionalität (neue Version) soll mit der Zeit die alte Funktionalität (alte Version) ersetzen, d.h. existierende Servicenutzer müssen nach einigen neuen Versionen (z.B. vier) auf die neue Version migrieren bevor der Service abgeschaltet werden kann. Dazu müssen sie den Migrationszeitraum abstimmen und Migrationsprojekte durchführen.
- Neue Servicevariante: Änderungen der Funktionalität des Services durch neue Anforderungen, die nur für einen Teil der Servicenutzer relevant bzw. akzeptabel sind. Die neue gilt dann als Variante des alten Services, ersetzt diesen also nicht. Die neue Servicevariante muss entwickelt, getestet, in Betrieb genommen und im Repository veröffentlicht werden.

Weitere Bestandteile dieser Framework-Komponente sind Richtlinien und Entscheidungsvorlage-Templates. Richtlinien sind Vorgaben für Verhaltensweisen und die Reihenfolge von Aktivitäten. Entscheidungsvorlage-Templates geben Aufschluss darüber, welche Entscheidung wann, von wem, wie schnell und unter Berücksichtigung welcher Informationen und Kriterien gefällt werden müssen.

3.5 Weitere Komponenten

Die weiteren Komponenten des Frameworks Zusammenarbeit und Kommunikation, Performance Management sowie Toolsupport (vgl. Abb. 4) werden hier aus Platzgründen nur kurz behandelt.

Mangelnde Zusammenarbeit und Kommunikation zwischen Fachabteilungen und IT ist ein häufiger Grund für Probleme. Durch regelmäßige Workshops, in denen gegenseitiges Feedback und Erfahrungen zu den Services ausgetauscht werden, können diese Probleme verhindert und die Effektivität der IT gesteigert werden.

Im Rahmen des Performance-Managements sind die Messung der Leistung der vorgegebenen Managementprozesse sowie deren Verbesserung, die Einhaltung der Richtlinien und die Messung der Zielerreichung der KPIs für die Rollen hervorzuheben. Aus CobiT können Messgrößen und Ziele für Prozesse und Rollen abgeleitet werden [IT07]. Weitere wichtige Aspekte des Performance-Managements, für die Designvorschläge erarbeitet wurden, sind die Klärung der Eigentumsfrage der Services und eine verursachungsge-rechte Kostenverrechnung.

Servicerepositories spielen für das Metadatenmanagement bezüglich der Services und anderer Artefakte sowie die darauf aufbauenden Managementprozesse eine wichtige Rolle. Sie erhöhen die Wirtschaftlichkeit und Akzeptanz des Servicemanagements. So werden im Rahmen des Frameworks Informationen zum Servicelebenszyklus, der Service-signatur und den Servicelevels aber auch zu Prozessen und anderen Artefakten im Repository abgelegt und gemanagt.

4 Evaluation in Workshops und Fallstudie

Das theoretische Literaturstudium war zunächst die Grundlage für das oben beschriebene Framework. Danach erfolgte verteilt über einen Zeitraum von vier Monaten die Evaluation in acht Workshops bei Firmen, die real eine SOA einführen bzw. dies planen. Zunächst wurden in einem Vorterminein die grundsätzliche Zielrichtung, Strategie und der Zeitrahmen der SOA-Einführung diskutiert sowie die individuellen Anforderungen erhoben. Das Autorenteam hat zudem mittels des oben kurz beschriebenen Fragebogens den Reifegrad bezüglich der Architektur und der IT-Governance im Unternehmen ermittelt. In jeweils einem ganztägigen anschließenden Workshop wurden nach individueller Vorbereitung dann jeweils geeignete Betriebsmodelle, Organisationsstrukturen, Rollen, Verantwortlichkeiten, Managementprozesse, Kostenverteilungskonzepte und Performancemetriken präsentiert, diskutiert und weiterentwickelt. Diskussionspartner waren SOA-Experten aus den zentralen Architekturmanagementabteilungen und wichtigen SOA-Projekten. Dabei stand die Anpassung der Konzepte an die spezifische Situation im jeweiligen Unternehmen im Vordergrund. Die diskutierten Anpassungen an den Konzepten wurden einerseits für das Unternehmen individuell aufbereitet und zur Verfügung gestellt. Andererseits wurden generell gültige Verbesserungen der Konzepte und deren Darstellung in das Ausgangsframework eingearbeitet und dann in den darauffolgenden Workshops als Ausgangsbasis verwendet.

Nach der ersten Phase der Evaluation und schrittweisen Verbesserung des Frameworks wurde die Evaluation mit einer Fallstudie fortgesetzt. Im Rahmen eines sechsmonatigen Projektes bei einem europäischen Energiedienstleister wurde begleitend zu einer SOA-Implementierung ein SOA-Governance-Konzept nach der vorgeschlagenen Methode und basierend auf dem Ausgangsmaterial entwickelt und zur Implementierung vorbereitet. Der Hauptfokus dieser Fallstudie lag auf der Entwicklung der grundlegenden Organisationsstruktur, des Rollenmodells und der Managementprozesse. Auch dieses Feedback wurde zur weiteren Verbesserung des Frameworks verwendet.

5 Fazit

In diesem Artikel wird ein Framework für das Management Serviceorientierter Architekturen vorgestellt. Es umfasst die vier IT-Governance-Komponenten Betriebsmodell und Organisationsstruktur, Managementprozesse und Richtlinien, Kommunikation und Zusammenarbeit sowie Performance Management und fokussiert inhaltlich auf das Unternehmensarchitekturmanagement und das SOA-Service-Management.

Anders als die bestehenden IT-Governance-Frameworks enthält es zahlreiche Anpassungen, um den SOA-spezifischen Herausforderungen gerecht zu werden. Im Gegensatz zu den bestehenden SOA-Governance-Ansätzen, die teilweise nur die besonderen Anforderungen dokumentieren und Lösungen nur auf ganz hoher Ebene oder sehr begrenzt betrachten, stellt dieses Framework einen zusammenhängenden und integrierten Ansatz zur Einführung der SOA-Governance dar. Insbesondere werden die verschiedenen Reifegrade der betrachteten Firma berücksichtigt und konkrete Designvorschläge für die praktisch evaluierten Problemfelder der SOA-Governance gegeben.

6 Literaturverzeichnis

- [Bi05] Bieberstein, N.; Bose, S.; Walker, L.; Lynch, A. (2005): Impact of service-oriented architecture on enterprise systems, organizational structures, and individuals. *IBM Systems Journal*. 44: 691-708.
- [Bi06] Bieberstein, N.; Bose, S.; Fiammante, M.; Jones, K.; Shah, R.: *Service Oriented Architecture Compass: Business Value, Planning, and Enterprise Roadmap*. IBM Press, 2006.
- [BI04] Bloomberg, J.: Zapthink (Hrsg.): *SOA Governance – IT Governance in the Context of Service Orientation*. http://www.logiclibrary.com/resources/zapthink_wp.php, 2004.
- [Bu05] Burschmid, M.: Vergleichende Buchbesprechung – IT-Governance. In: *Wirtschaftsinformatik* 47 (2005) Nr. 6; S. 448-463.
- [Cm06] CMMI Product Team: *CMMI for Development – Version 1.2*; Pittsburgh, PA: Carnegie Mellon University, Software Engineering Institute 2006, <http://www.sei.cmu.edu/publications/documents/06.reports/06tr008.html>.
- [DH08] Dico, A.; Hornford, D. (2008): *Delivering SOA with TOGAF*.
- [Gr03] van Grembergen, W.: *Strategies for information technology governance*. Idea Group Publishing, London, 2003.
- [He07] Heutschi, R.: *Serviceorientierte Architektur – Architekturprinzipien und Umsetzung in die Praxis*. Springer, Berlin, 2007.
- [IT03] IT Governance Institute: *Governance implementation guide: How do I use CobiT to implement IT governance?* IT Governance Institute, Rolling Meadows, 2003.
- [IT07] IT Governance Institute: *CobiT 4.1 – Framework, Control Objectives, Management Guidelines, Maturity Models*. IT Governance Institute, Rolling Meadows, IL, 2007.

- [KBS05] Krafzig, D.; Banke, K.; Slama, D.: Enterprise SOA – Service-Oriented Architecture Best Practices. Prentice Hall Professional Technical Reference, Upper Saddle River, New Jersey, 2005.
- [KS07] Kastner, P. S.; Saia, R.: Management and governance: Planning for an optimized SOA application lifecycle; Boston: Aberdeen Group, Inc. 2007, http://www.aberdeen.com/summary/report/benchmark/3944_RA_SOAGov.asp.
- [MB06] Marks, E. A.; Bell, M.: Executive's Guide to Service-Oriented Architecture. John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.
- [MN03] McCoy, D.; Natis, Y. (2003): Service-Oriented Architecture: Mainstream Straight Ahead. Gartner Research.
- [OGC07] Office of Government Commerce (OGC): ITIL Lifecycle Publication Suite Books. TSO (The Stationary Office), 2007.
- [Pe04] Peterson, R.: Crafting Information Technology Governance. In: Information Systems Management Journal Fall (2004).
- [PMB05] Plessis, J. J. d.; McConville, J.; Bagaric, M.: Principles of contemporary corporate governance. Cambridge University Press, Cambridge, 2005.
- [RW04] Ross, J. W.; Weill, P.: IT Governance: How Top Performers Manage It Decision Rights for Superior Results. 2004.
- [RWR06] Ross, J. W.; Weill, P.; Robertson, D. C.: Enterprise Architecture as Strategy. Harvard Business School Press, Boston, MA, 2006.
- [SS06] Schelp, J.; Stutz, M.: SOA-Governance. In: HMD – Praxis der Wirtschaftsinformatik (2006) Nr. 253; S. 66-73.
- [SS07] Simonsson, M.; Schröpfer, C.: Setting the IT organization goals. In: Johnson, P., Ekedt, M. (Hrsg.): Enterprise Architecture – Models and analysis for information systems decision making. Studentlitteratur, 2007; S. 153-212.
- [ST07] Starke, G.; Tilkov, S. (Hrsg). SOA-Expertenwissen – Methoden, Konzepte und Praxis serviceorientierter Architekturen. Heidelberg 2007: dpunkt.verlag.
- [Sy06] Systinet: SOA Governance: Balancing Flexibility and Control Within an SOA. http://www.systinet.com/dl/SOA_Gov906.pdf, 2006.
- [Wo03] Woods, D.: Enterprise Service Architecture. O'Reilly & Associates, Cambridge, 2003.
- [Wo06] Woolf, B.: Introduction to SOA Governance – Governance: The official IBM definition, and why you need it. <http://www-128.ibm.com/developerworks/webservices/library/ar-servgov/index.html>, 2006.

Integration öffentlicher Anwendungslandschaften

Norbert Gronau, Moreen Stein, Sandy Eggert, Anne Lämmer

Lehrstuhl für Wirtschaftsinformatik und Electronic Government

Universität Potsdam

August-Bebel-Straße 89

14482 Potsdam

{n Gronau | m Stein | s Eggert} @ wi.uni-potsdam.de

anne.laemmer@sdm.de

Abstract: Öffentliche Anwendungslandschaften sind durch zahlreiche Individuallösungen gekennzeichnet, die jeweils eine einzelne Funktion im Leistungserstellungsprozess abbilden. Darüber hinaus sind die Anwendungen, wenn überhaupt, nur bilateral gekoppelt. Dies führt zu hohen Administrations- und Wartungskosten und schränkt ferner die Wandlungsfähigkeit der Architekturen ein. Zwar sind Teilprozesse automatisiert, es existiert jedoch keine durchgehende Integration entlang der Verwaltungsabläufe. Anhand eines Fallbeispiels werden ein Vorgehensmodell und Lösungsansätze zur Integration von öffentlichen Software-Anwendungen dargestellt.

1 Einleitung

Die Anwendungslandschaften öffentlicher Verwaltungen sind im Vergleich zu Unternehmen durch zahlreiche Individuallösungen gekennzeichnet. Öffentliche Anwendungslandschaften umfassen sämtliche Fachanwendungen, die im öffentlichen Sektor betrieben werden, und ihre Beziehungen untereinander. Infrastrukturelemente wie Server oder Netzwerkverbindungen finden keine Berücksichtigung. Die Ursachen für den hohen Anteil an meist langjährig genutzten Individuallösungen liegen in der weit verbreiteten Ansicht, dass sich Verwaltungen und Unternehmen grundlegend voneinander unterscheiden. Dies betreffe sowohl die typischen Geschäftsprozesse, wie Antragsbearbeitung oder Kontrollen, als auch die Leitgedanken des organisationalen Handelns, die sich aus dem Vollzug des übergeordneten politischen Willens ableiten. An dieser Stelle zeigen sich die Prinzipien Regelgebundenheit, Aktenmäßigkeit, Hierarchie und klare Zuständigkeit [Ma94]. Durch die Betonung dieser Unterschiede wurden besonders in der Vergangenheit individuelle Lösungen, die seltene Spezialfälle bedienen, Standardlösungen vorgezogen. Gleichwohl führt unter anderem das Messen an privaten Maßstäben zu einem erhöhten Veränderungsdruck [BAF07]. Dies betrifft vor allem die Gestaltung der internen Prozesse, deren Teilprozesse oftmals nur isoliert durch Individual- und Speziallösungen abgebildet werden. Auch wenn es vermehrt gelingt, die Schnittstelle zu Unternehmen und Bürgern zu verbessern, so zeigen sich die Schwächen vor allem durch die Dauer des Prozesses der internen Bearbeitung, die auf der Ebene Government-to-

Government positioniert ist. Es stellt sich die Frage, wie die Anwendungslandschaft von öffentlichen Verwaltungen in der Weise gestaltet werden kann, dass auch grenzüberschreitende Prozesse medienbruchfrei ablaufen können.

Der vorliegende Artikel beschreibt zunächst typische Anwendungslandschaften in öffentlichen Verwaltungen, anschließend wird ein Vorgehen zur Erhebung und Analyse dieser Landschaften sowie zur bedarfsgerechten Integration von E-Government-Anwendungen aufgezeigt. Die Ergebnisse wurden im Rahmen eines Forschungsprojektes gemeinsam einem deutschen Bundesland gewonnen und erprobt. Ziel war es, eine Integrationslösung für die heterogene Anwendungslandschaft des Praxispartners zu entwerfen, die schrittweise umsetzbar ist.

2 Charakteristika und Integrationsbedarf

Öffentliche Anwendungslandschaften sind durch viele Einzelanwendungen gekennzeichnet, die jeweils einer spezifischen Funktion dienen und in vielen Fällen ausschließlich innerhalb einer Behörde genutzt werden. Diese Vielfalt und der isolierte Einsatz basieren auf der Breite des Aufgabenspektrums [Ho05] und auf dem hohen Spezialisierungsgrad der öffentlichen Verwaltungen [Wi06]. Beides findet sich sowohl in der Organisationsstruktur als auch in der Anwendungslandschaft wieder. Die Aufgaben einzelner Fachbereiche werden jeweils durch eine Vielzahl von Anwendungen abgebildet [Du00], wobei oftmals Individualsoftware eingesetzt wird [Ge00]. Gleichwohl ist es notwendig, dass Verwaltungsleistungen behördenübergreifend erstellt werden können, hierzu ist eine entsprechende Integration zwischen den Anwendungen grundlegend. Vermehrt bestehen auf Landesebene schon Integrationsmechanismen durch Querschnittsanwendungen wie behördenübergreifend genutzte Intranetportale oder ERP-Systeme. Es fehlt jedoch an umfassendem und gesichertem Wissen über die Gesamtarchitektur und somit auch an einer bedarfsgerechten Integration, die mit dem Aufbau von Enterprise Architecture Management einhergehen muss. Vorhaben mit dem Ziel der Integration verlangen gründliche Kenntnisse der eigenen Anwendungslandschaft [Ge00]. Es ist daher ein Verfahren zu entwickeln, das regelmäßig die landesweite E-Government-Landschaft erhebt und auswertet. Dies erfordert zunächst einen hohen Koordinationsaufwand. Neben den informationstechnischen Architekturelementen müssen auch die ablaufenden Prozesse zugeordnet werden. Nur so kann der Integrationsbedarf identifiziert werden. Dies ist in Zusammenarbeit mit den einzelnen Behörden durchzuführen. Die Integration muss sich an fachlichen Anforderungen unter Berücksichtigung der Verwaltungsprozesse orientieren. Weiter ist eine zentrale Steuerungseinheit für die Datenerhebung im Sinne eines landesweiten Architekturmanagements für die Qualitätssicherung der Daten grundlegend. Nur durch eine breite Beteiligung können belastbare Ergebnisse abgeleitet werden, die nach einer Analyse zu neuen konkreten Integrationsprojekten führen. Ferner sollten auch die in vertikaler Richtung angrenzenden Behörden einbezogen werden, das heißt im Fall eines Landes kommunale und Bundesbehörden. Auch hier sind die entsprechenden Prozesse und die dazugehörigen Architekturelemente aufzunehmen. Es können Tools zum umfassenden Architekturmanagement Verwendung finden, welche die aktuelle und zukünftige Anwendungslandschaft abbilden [Er06].

3 Analyse öffentlicher Anwendungslandschaften und Lösungsansatz zur schrittweisen Integration

Das Vorgehen zur Analyse öffentlicher Anwendungslandschaften orientiert sich an den drei Schritten Erhebung, Modellierung und Auswertung. Im Fall des Praxispartners wurden fünf Szenarien ausgewählt, für die jeweils die Integrationsbedarfe vor dem Hintergrund der Eigenschaften der Gesamtlandschaft ermittelt wurden. Es ist kaum umsetzbar, die gesamte Anwendungslandschaft einer Landesverwaltung zu erheben, eine Priorisierung fördert hier den Realisierungserfolg [Ho05]. Die Auswahl der Szenarien sollte sich daher an vorab definierten Kriterien orientieren, die im Folgenden vorgestellt werden. Auf Basis der Modellierung und Auswertung wird ein Lösungsansatz zur schrittweisen Integration vorgestellt. Die Darstellungen beschränken sich auf die Anwendungsebene und berücksichtigen keine Infrastrukturelemente, wie zum Beispiel Netzwerkkomponenten.

3.1 Erhebung der Anwendungen

Zur Erhebung der Anwendungslandschaft sind zunächst die bestehenden Dokumente auf ihre Eignung zur Datenaufnahme zu prüfen. In der ersten Phase der Erhebung stehen die Identifikation sämtlicher Fachanwendungen sowie deren Schnittstellen untereinander im Zentrum. Eine Fachanwendung implementiert Geschäftskomponenten (Prozessschritte, z.B. Antragsbearbeitung), beinhaltet Softwarekomponenten und nutzt Infrastrukturkomponenten (z.B. Server, Internet/Intranet) [Ni05]. Genügen die existierenden Datenbestände zu den bestehenden Fachanwendungen nicht, ist eine ergänzende schriftliche Befragung durchzuführen. Diese kann sich an die Administratoren oder IT-Verantwortlichen richten. Grundsätzlich wären auch persönliche Befragungen denkbar, da eine Landesverwaltung jedoch aus einer Vielzahl von verschiedenen Behörden mit jeweils eigenen Anwendungen besteht, ist eine schriftliche Befragung praktikabler. Die so gewonnen Daten bilden die Grundlage für die Analyse der Gesamtarchitektur. Die unter Abschnitt 2 beschriebenen typischen Eigenschaften von öffentlichen Anwendungslandschaften werden durch die Ergebnisse der Erhebungen bei dem Praxispartner bestätigt. Ein Großteil der voneinander isolierten Fachanwendungen wird nur durch wenige Mitarbeiter genutzt. Relevant sind in dieser Stufe neben der Identifikation von Fachanwendungen und deren Schnittstellen auch die Zuordnung zu der Organisationseinheit sowie die fachliche Einordnung in einen Prozess oder ein Tätigkeitsfeld. Dies erleichtert den anschließenden Wechsel auf die Prozesssicht, da in der zweiten Phase der Datenerhebung die Fachverantwortlichen adressiert werden. Sie kennen die ablaufenden Prozesse und die verwendeten Anwendungen und können bereits eine erste Einschätzung zum Integrationsbedarf abgeben. Dafür sind je identifizierter Fachanwendung die nutzenden Prozesse abzufragen. Die Prozessbetrachtung bildet die sichere Basis für Auswahl geeigneter Integrationslösungen [Sc03]. Weiter müssen manuell organisierte Schnittstellen zu anderen Anwendungen oder Bearbeitern erfasst werden. Die Integrationsrelevanz lässt sich außerdem über die Anzahl der nutzenden Behörden und Mitarbeiter, die Frequenz des Prozessablaufs sowie die geschätzten Verbesserungspotenziale, bezogen auf Zeit-, Kosten- und Qualitätsvorteile, einschätzen. Je höher sich die Integrationsrelevanz

je aufgenommenen Prozess zeigt, umso eher sollte auf Basis dieses Prozesses das entsprechende Integrationsszenario gebildet werden.

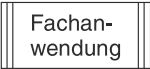

Phase	Sicht	Erhebungsmethode	Kriterien
1	 Fachanwendung	Schriftliche Befragungen	<ul style="list-style-type: none">▸ Nutzende Mitarbeiter/Behörden▸ Schnittstellen
2	 Prozess	Telefoninterviews	<ul style="list-style-type: none">▸ Einschätzung der Integrationsrelevanz▸ Fallzahlen▸ Geschätzte Verbesserungspotenziale

Abbildung 1: Ermittlung der Integrationsrelevanz

Das Integrationsszenario umfasst einen abgrenzbaren Prozess, der verschiedene Fachanwendungen nutzt. Ein Beispiel aus dem Anwendungsfall ist der Prozess Neueinstellung in den Schuldienst, der verschiedene Fachanwendungen tangiert. Das Szenario beinhaltet folglich diesen Prozess sowie verschiedene Rahmenbedingungen (zum Beispiel gesetzliche Anforderungen bezogen auf die Form der eingehenden Bewerbungen oder geplante IT-Investitionen). Nicht immer sind sämtliche erforderlichen Daten zur Bewertung des Integrationsbedarfs in angemessener Qualität verfügbar. Dies kann zum einen an einer möglichen fehlenden Kooperationsbereitschaft der betroffenen Verwaltungseinheiten liegen, zum anderen aber auch auf eventuell auftretende Fehleinschätzungen und -aussagen durch die Befragten zurückzuführen sein. Es ist daher nicht sicher feststellbar, ob die mit dem hier beschriebenen Vorgehen ausgewählten Szenarien tatsächlich den höchsten oder lediglich einen beachtlichen Integrationsbedarf aufweisen. Gleichwohl hat sich das Verfahren insgesamt bei der Erprobung als praktikabel erwiesen. Zwar wäre es auch denkbar, zunächst sämtliche Prozesse zu erfragen. Dies führt jedoch dazu, dass bereits erste Hürden zum Verständnis des Prozessbegriffs auftauchen. Ferner ist es weitaus aufwändiger je Verwaltungseinheit die relevanten Prozesse zu identifizieren, aufzunehmen und die genutzten Anwendungen zuzuordnen als über eine meist existierende erste Erhebung der genutzten Softwareanwendungen, die relevanten Fachanwendungen zu identifizieren. Die IT-Verantwortlichen fungieren hier als erster, meist auch einheitlicher Ansprechpartner und können bereits Angaben zu den existierenden und geplanten Schnittstellen liefern.

3.2 Modellierung der Anwendungslandschaft

Die erhobenen Daten ermöglichen die Visualisierung der bestehenden Anwendungslandschaft. Zu diesem Zweck wird die Softwarekartographie eingesetzt. Diese beschäftigt sich mit der Darstellung von IT-Landschaften durch Softwarekarten und dient neben einem besseren Überblick auch der besseren Beherrschbarkeit der hohen Komplexität dieser Anwendungslandschaft und dem Erkennen von Veränderungen der Anwendungslandschaft. Die Softwarekarten dienen der Repräsentation der Anwendungslandschaft mit dem Fokus auf der Gestaltung und Planung der komplexen Informationsinfrastruktur. Eine Softwarekarte setzt sich aus einer oder mehrerer Sichten, die verschiedene

Aspekte visualisieren, zusammen [MW04]. Diese einzelnen Sichten bestehen aus einem Kartengrund und aufbauenden Schichten, welche verschiedene Informationen bezüglich der Anwendungslandschaft transportieren.

Zur Umsetzung der kartographischen Darstellung empfiehlt sich in Integrationsprojekten unter der Berücksichtigung der Anforderungen die Nutzung von Cluster- und Prozesskarten. Clusterkarten erlauben es, Informationssysteme logischen Einheiten zuzuordnen. Hierbei gibt der Kartengrund eine Clusterung vor. In Schichten lassen sich neue Cluster und Anwendungen aufbringen, die bei Bedarf miteinander verbunden werden können, um sie in Beziehung zu setzen [Er06]. Eine Clusterkarte lässt sich somit durch Kategorien wie zum Beispiel Funktionsbereiche, Organisationseinheiten, Standorte, Regionen usw. einteilen. Weiterhin können Konventionen festgelegt werden, die nach bestimmten Kriterien die Position der Cluster festlegen, um die Lesbarkeit der Karten weiter zu erhöhen [La05].

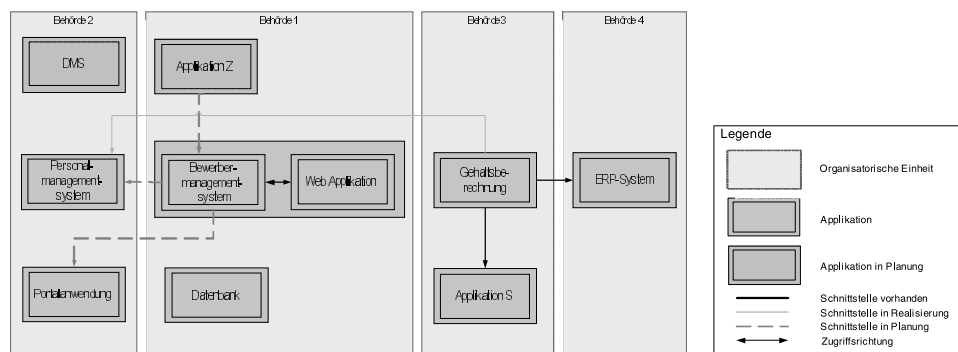


Abbildung 2: Beispiel für eine Clusterkarte

Darüber hinaus können Prozesskarten eingesetzt werden, die die Anwendungslandschaft der aufgenommenen Prozesse darstellen. Hierbei werden auf der horizontalen Dimension Prozesse bzw. Prozessschritte aufgebracht, während das zu visualisierende Merkmal, bzw. Entitäten denen Anwendungssysteme zugeordnet werden sollen, die vertikale Dimension bilden [La05]. Die Verortung eines Anwendungssystems transportiert die Information, welche Prozesse das Anwendungssystem abbildet und welche Ausprägung das visualisierte Merkmal annimmt, bzw. welcher Entität des auf der Vertikalen visualisierten Typs das Anwendungssystem zugeordnet ist. Welcher Natur diese Beziehung ist, muss in einer Legende festgelegt werden [LMW05a]. Diese Darstellung erlaubt es beispielsweise, Anwendungen zu identifizieren, die bereichsübergreifend Verwendung finden, oder bei der Durchführung mehrere Prozessschritte benötigt werden und dient daher der Darstellung von Integrationsbedarfen. Prozesskarten visualisieren die fachlichen Aspekte der Anwendungslandschaft.

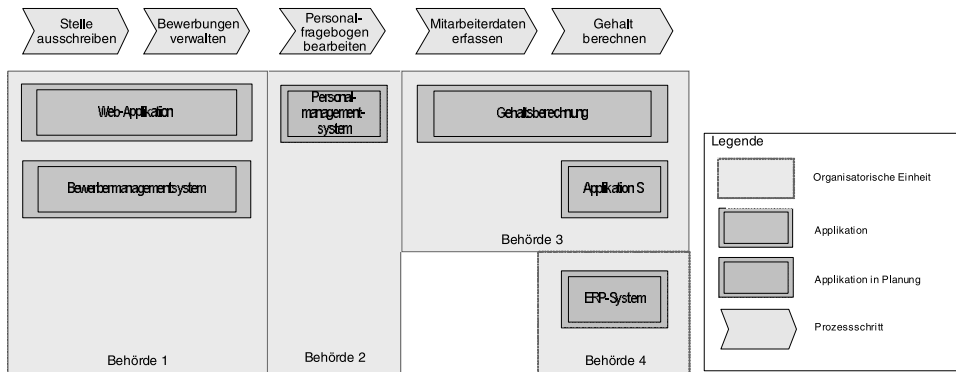


Abbildung 3: Beispiel für eine Prozesskarte

Die Softwarekartographie orientiert sich in ihrer Darstellung und den Methoden an herkömmlichen Landkarten. Dadurch gewährleisten sie einen einfachen Zugang zur Darstellung von Anwendungslandschaften [LMW05b]. Dies ist besonders zur Begründung und Erläuterung von Integrationsvorhaben gegenüber betroffenen Verwaltungseinheiten von Vorteil. Integrationsbedarfe können anhand der vorgestellten Prozesskarten leicht abgelesen und kommuniziert werden. Auf Basis dieses gemeinsamen Verständnisses lassen sich auch Anforderungen aus fachlicher Sicht verbessert diskutieren und in den Lösungsansatz zur Integration aufnehmen.

3.3 Auswertung und Lösungsansatz

Im Fall des Praxispartners hat sich gezeigt, dass die Einführung einer einheitlichen Integrationslösung nicht sinnvoll ist. Zum einen existieren fachbezogene Insellösungen, bei denen zunächst kein Integrationsbedarf besteht, zum anderen erscheinen je nach Einzelfall andere Integrationsansätze sinnvoll. Weiter ist es kaum möglich, innerhalb eines Integrationsprojektes die gesamte Anwendungslandschaft einer Landesverwaltung zu integrieren. Vielmehr vollzieht sich die Integration schrittweise. Daher ist es günstig, lokale Lösungen einzuführen und diese sukzessiv und vor allem bedarfsorientiert zu integrieren. Insgesamt sollten jedoch Regeln für die Ausgestaltung der Anwendungslandschaft existieren, um standardisierte Integrationsmuster zu erzielen.

Sind sämtliche Fachanwendungen mit ihren automatisierten wie auch manuellen Schnittstellen erhoben, ergibt sich in der grafischen Darstellung ein Netz aus Knoten (Anwendungen) und Kanten (Schnittstellen). Dies kann mittels einer Clusterkarte erfolgen. Dieses Netz wird an einigen Stellen dichter sein als an anderen. Es zeigen sich Wolken, die eine Domäne abbilden. Überall dort, wo das Netz aus Anwendungen und Schnittstellen besonders dicht ist, tauschen mehrere Anwendungen Informationen aus. Geschieht dies regelmäßig in vielen Fällen (darstellbar durch die Stärke der Kanten und Knoten), wird eine Domäne gebildet. Der Kartengrund wird durch die zugehörigen Behörden definiert, was die Bildung der Domänen erleichtert. Nicht immer genügt die rein mathematisch orientierte Auswertung des Anwendungsnetzes. Die betroffene Behörde ist ebenfalls zu berücksichtigen. Zum Beispiel ist es aus rechtlichen Gründen nicht immer möglich, dass

verschiedene Behörden bestimmte, personenbezogene Daten austauschen. Aier beschreibt ein Verfahren, das Domänen (benannt als Cluster) für die Bereitstellung von Diensten einer Serviceorientierten Architektur (SOA) durch die aus den Sozialwissenschaften stammende Netzwerkanalyse bestimmt [Ai06]. Grundlage für diese Berechnungen sind die erhobenen Geschäftsprozesse, die jeweils genutzten Anwendungen sowie deren Schnittstellen. Das Verfahren bietet erste Hinweise zur Definition von Services. Es berücksichtigt jedoch keine zukünftigen Anforderungen, die sich bisher nicht in der Architekturlandschaft widerspiegeln. Auch Aspekte der Zielarchitektur müssen Beachtung finden. Entscheidend sind die logische Zusammengehörigkeit der Fachanwendungen bezüglich der Geschäftsprozesse und ihre Abhängigkeit untereinander. Thematisch verwandte Anwendungen werden gruppiert und bilden eine Domäne, innerhalb derer die Anwendungen integriert werden und Daten bezogen auf verschiedene Prozesse austauschen. Ein Beispiel aus dem Praxisfall ist die Domäne kommunale Steuern (zum Beispiel Gewerbesteuern). Sämtliche Prozesse, die bezogen auf kommunale Steuern in Landesbehörden ablaufen, können über eine entsprechende Integration der Anwendungen durch die Domäne abgebildet werden.

Innerhalb einer Domäne kann jeweils ein anderer Integrationsansatz gewählt werden, um die Informationssysteme, die zur Bearbeitung eines Fachverfahrens benötigt werden, zu integrieren. Dabei ist jede Technik denkbar, sei es als bereits bestehende Punkt-zu-Punkt-Verbindung, als Hub-and-Spoke-Architektur oder als eine Serviceorientierte Architektur. Die Domänen sind nach außen abgeschottet und stellen ihre Leistungen im Bedarfsfall anderen Domänen über eine globale Integrationslösung zur Verfügung. Das heißt, dass keine individuellen Schnittstellen (manuell oder automatisiert) zwischen Anwendungen aus einer Domäne und Anwendungen außerhalb der Domäne existieren sollten. Ist ein Datenaustausch notwendig, erfolgt dieser über eine einheitliche, globale Lösung, durch die sämtliche Domänen integriert werden können. Wer die Domänen und die damit verbundene Integrationslösung fachlich und administrativ betreut, richtet sich nach den jeweiligen Rahmenbedingungen. Sind eindeutige fachliche Zuständigkeiten gefordert, eignet sich eher die Betreuung durch eine Behörde. Ein fachlich unabhängiger Betrieb (zum Beispiel durch eine Datenzentrale) ist dann sinnvoll, wenn das technische Wissen in den Behörden fehlt und die fachliche Betreuung nicht permanent notwendig ist.

Allerdings genügt die Aufspaltung der Organisation in Domänen noch nicht, um sämtliche möglicherweise notwendigen Arten der Integration zu beschreiben. Es existieren ferner Querschnittsanwendungen, die durch ihren Funktionsumfang einen integrierenden Charakter besitzen. Hier wird in der Regel keine Integrationslösung benötigt, vielmehr müssen die notwendigen Fachverfahren und Behörden Zugang zu diesen Systemen erhalten. Typisch hierfür sind zentrale Verfahren bezüglich Personal und Haushalt, es eignen sich aber auch Geodatenportale, Archivierungsanwendungen oder Kollaborationswerkzeuge als Querschnittsanwendungen. Ähnlich wie die Querschnittsanwendungen können von zentraler Stelle auch Shared Services für standardisierte Funktionen in die Architektur eingebunden werden. Zum Beispiel eignen sich E-Payment-Systeme oder Formularserver als Shared Services. Der Unterschied zwischen Shared Services und Querschnittsanwendungen liegt in dem verschieden ausgeprägten Funktionsumfang. Shared Services bilden standardisierte, häufig genutzte Prozessschritte ab [JW04]. Die

vorhandenen Redundanzen, zum Beispiel verschiedene E-Payment-Systeme je Verwaltungseinheit, werden auf diese Weise abgebaut [Li07]. Querschnittsanwendungen, zum Beispiel ERP-Systeme, umfassen weitaus mehr Funktionen und bilden nicht nur einzelne Prozessschritte ab.

Das hier vorgestellte Integrationsmodell steigert die Wandlungsfähigkeit der Anwendungslandschaft. Wandlungsfähigkeit beschreibt die Fähigkeit eines Systems, sich selbst effizient und schnell an veränderte Anforderungen anpassen zu können und wird mittels verschiedener Indikatoren (zum Beispiel Modularität, Skalierbarkeit, Interoperabilität, Selbstähnlichkeit) messbar gemacht [An06, Gr06]. Vor dem Hintergrund sich häufig ändernder rechtlicher Rahmenbedingungen, die sich sowohl auf die Prozesse als auch auf die Organisationsstruktur öffentlicher Verwaltungen auswirken, stellt Wandlungsfähigkeit eine grundlegende Anforderung an öffentliche Anwendungslandschaften dar. Die beschriebenen Domänen bilden klar abgegrenzte Einheiten, auch beschreibbar als Module. Auf diese Weise wird das Hinzufügen und Entfernen von Domänen erleichtert, was wiederum die Skalierbarkeit der Gesamtarchitektur positiv beeinflusst. Weitere Vorteile dieses Architekturmodells ergeben sich durch die bedarfsorientierte Clusterintegration. So muss nicht jede Fachanwendung mit jeder anderen verbunden werden. Eine durchgängige Serviceorientierte Architektur ist zwar bezogen auf Wandlungsfähigkeit erstrebenswert, deren Implementierung ist jedoch sehr ressourcenaufwändig. Die Integration innerhalb der Domänen kann schrittweise erfolgen. Die sukzessive Implementierung wird durch eine geringere Anzahl der verantwortlichen wie auch betroffenen Mitarbeiter erleichtert. Die Komplexität der Gesamtlandschaft wird in kleinere Landschaften überführt und so handhabbar. Es entstehen realisierbare lokale Integrationslösungen, die in ihrem Zusammenspiel als Serviceorientierte Architektur eine lose gekoppelte Anwendungslandschaft bilden.

4 Kriterien zur Auswahl der Integrationslösung

Nachdem die Domänen identifiziert sind, stellt sich die Frage, welche Integrationsart innerhalb der Domänen geeignet ist. Dafür werden im Folgenden zunächst die Integrationsarchitekturen Punkt-zu-Punkt, Hub-and-Spoke und SOA bezogen auf mögliche Rahmenbedingungen eingeordnet. Darüber hinaus finden auch Querschnittsanwendungen und Shared Services Berücksichtigung. Anschließend folgt eine von den Strategien Punkt-zu-Punkt, Hub-and-Spoke und SOA unabhängige Kategorisierung der Integrationsarten, die eine verstärkt fachlich orientierte Auswahl ermöglicht.

Eine Punkt-zu-Punkt-Architektur (PzP-Architektur) ist dann sinnvoll, wenn diese bereits besteht, d.h. alle Fachanwendungen einer Domäne sind bereits miteinander verbunden und es besteht kein weiterer Integrationsbedarf. In diesem Falle ist es möglich, diese PzP-Architektur solange beizubehalten, bis sich ein Integrationsbedarf einstellt. Die Neueinrichtung oder Weiterentwicklung einer PzP-Architektur empfiehlt sich jedoch nicht, auch wenn diese bei einer ersten Kosten/Nutzen-Rechnung günstiger erscheinen mag [Vo05]. Bezogen auf die Zukunftsfähigkeit hat diese Architektur keine Erfolgsaussichten, da sie kaum wandelbar ist und sich so nur schlecht an neue Rahmenbedingungen anpassen lässt.

Ausschlaggebend für eine Entscheidung zwischen Hub-and-Spoke und SOA ist die adressierte Integrationsebene. Bei einer Integration auf der Datenebene ist eher eine datenorientierte Middleware oder auch ein Portalansatz, bei dem über ein Portal mehrere Datenquellen angefragt werden können, einzusetzen. Eine Integration auf der Applikations- oder Objektebene stellt den klassischen Hub-and-Spoke-Anwendungsbereich dar; busorientierte Middleware ist hier ebenso geeignet wie Applikationsserver oder Messagingssysteme. Bei der Integration der Prozessebene ist eine SOA als Integrationsarchitektur anzustreben. Die Geschäftsprozesse werden dabei unabhängig von den operativen Systemen betrachtet und beschreiben damit einen Workflow, der die entsprechenden Dienste der einzelnen Anwendungen aufruft [Sc05]. Mit serviceorientierten Architekturen können heterogene Elemente lose miteinander verknüpft werden, wodurch die Integration bezogen auf übergreifende Geschäftsprozesse verbessert werden kann [TF07].

Das Integrationsobjekt dient ebenfalls der Entscheidungsfindung. Sollen beispielsweise mehrere Anwendungen auf die gleichen Daten zugreifen, dann empfiehlt sich ein Hub-and-Spoke. Kann eine zentrale Anwendung, auf die von mehreren anderen Anwendungen zugegriffen werden muss, identifiziert werden, eignet sich ebenfalls ein Hub-and-Spoke. Werden in einem Prozess jedoch mehrere Anwendungen benötigt, ohne dass eine davon als zentral bezeichnet werden kann, dann sollte die Anbindung über eine SOA in Betracht gezogen werden. Werden ferner Verwaltungsaufgaben, die in allen Behörden anfallen und gleichartig gestaltet sind, in einem einzigen operativen Anwendungssystem abgebildet, stellt dieses System eine Querschnittsanwendung dar. Treten in verschiedenen Prozessen und in verschiedenen Domänen zusätzlich zu den Fachanwendungen auch gleiche Anwendungssysteme auf, so sollte eine genauere Untersuchung auf deren Eignung zu einem Shared Service durchgeführt werden. Weitere Hinweise und Kennzeichen für die Entscheidung sind in Abb. 4 dargestellt.

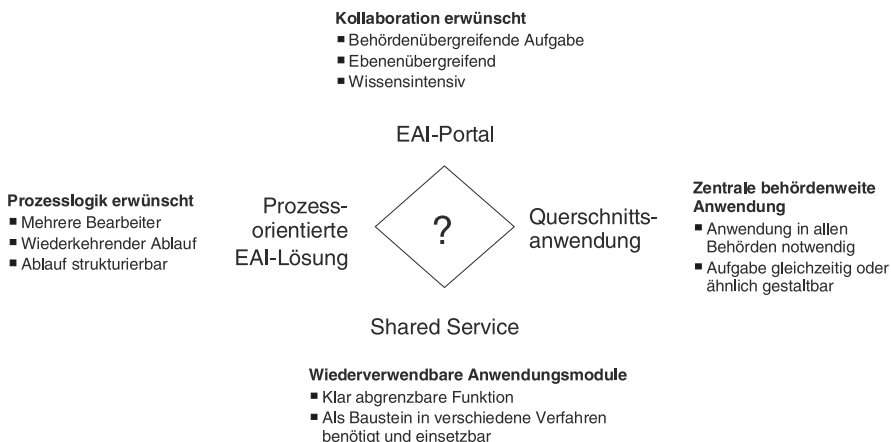


Abbildung 4: Hinweise und Kennzeichen für die Auswahl einer Integrationslösung

Eine Unterscheidung in Punkt-zu-Punkt, Hub-and-Spoke oder SOA wird hier nicht vorgenommen. Am Markt als EAI-Lösung (Enterprise Application Integration) angebotene

Integrationslösungen, sind in der Lage, unterschiedliche Integrationstechniken umzusetzen. Relevant ist hier die Frage, ob eine prozessorientierte Lösung, eine Portallösung oder eine Mischlösung als Integrationssystem geeignet ist.

5 Zusammenfassung und Ausblick

Dieser Beitrag beschreibt ein domänenbasiertes Architekturmodell, das ein sukzessives und bedarfsgerechtes Vorgehen zur Herausbildung einer integrierten Anwendungslandschaft ermöglicht. Ausgehend von der Aufnahme und Analyse der Unternehmensarchitektur werden mit Hilfe der Softwarekartografie Domänen identifiziert und nach der Erhebung des Integrationsbedarfs ein Integrationsansatz für die jeweilige Domäne entwickelt. Eine Domäne umfasst mehrere Anwendungen die fachlich einen engen Bezug zueinander haben und über Schnittstellen zueinander verfügen müssen. Domänen orientieren sich nicht an den Grenzen von Organisationseinheiten. Innerhalb einer Domäne können neben dem Aufbau einer serviceorientierten Architektur auch andere Integrationsansätze, wie der Aufbau einer Hub-and-Spoke-Architektur, durch verschiedene Typen von EAI-Lösungen angewandt werden. Neben den Domänen existieren Querschnittsanwendungen, die bereits aufgrund ihrer jeweiligen Aufgabe eine integrierende Wirkung besitzen, dies kann zum Beispiel ein ERP-System sein. Querschnittsanwendungen stellen keine eigene Domäne dar, zeichnen sich aber dadurch aus, dass der Großteil der Fachabteilungen einen Zugriff auf diese Systeme benötigt. Weitere, stark von standardisierten Vorgängen geprägte, Anwendungen können als Shared Services identifiziert werden. Die domänenorientierte Integration fördert die Umsetzbarkeit des Integrationsvorhabens, da nicht die gesamte Anwendungslandschaft innerhalb eines Projektes integriert wird, sondern ein schrittweises Vorgehen erfolgen kann.

Die Identifikation der Domänen durch Methoden der Softwarekartografie konnte anhand der praktischen Erprobung gemeinsam mit dem Bundesland in einem prototypischen Rahmen umgesetzt werden. Der nächste Schritt ist, diese Erkenntnisse, besonders bezogen auf die Bildung der Domänen und die Auswahl der Integrationsart, in weiteren Fällen zu prüfen und zu verfeinern. Ziel ist die Erstellung eines Kriterienkataloges, der es ermöglicht, die Ausgestaltung der Integrationslösung entsprechend der jeweiligen Rahmenbedingungen abzuleiten. Ferner muss auch die vertikale Integration verstärkt eingebunden werden. Zwar wurden im Forschungsprojekt mit dem Bundesland auch vereinzelte Beispiele mit Schnittstellen zu kommunalen oder Bundesbehörden herangezogen, sichere Erkenntnisse zur Integration können jedoch erst nach weiteren Untersuchungen gewonnen werden. Schließlich müssen Verwaltungen in der Lage sein, ihre Anwendungslandschaft nicht nur im Rahmen eines einzelnen Projektes zu verbessern, sondern diese auch in langfristiger Perspektive gestalten zu können. Methoden und Werkzeuge des Enterprise Architecture Managements etablieren sich in Unternehmen und bilden auch für öffentliche Verwaltungen einen geeigneten Weg, ihre Anwendungslandschaften bedarfsgerecht zu formen. Das hier beschriebene Vorgehen wurde im Kontext öffentlicher Verwaltungen entwickelt und erprobt, es eignet sich bei ähnlichen Ausgangsbedingungen (unter anderem zahlreiche Individuallösungen, hoher Spezialisierungsgrad) jedoch auch für Integrationsvorhaben in Unternehmen des privaten Sektors.

6 Literaturverzeichnis

- [Ai06] Aier, S.: How Clustering Enterprise Architectures helps to Design Service Oriented Architectures. In: Proceedings of the IEEE International Conference on Services Computing (SCC'06), Chicago 2006. IEEE Computer Society, Los Alamitos, 2006; S. 269–272.
- [An06] Andresen, K.: Design and Use Patterns of Adaptability in Enterprise Systems. GITO-Verlag, Berlin, 2006.
- [BAF07] Becker J; Algermissen, L.; Falk, T.: Prozessorientierte Verwaltungsmodernisierung. Springer, Berlin, Heidelberg, 2007.
- [Du00] Dubbert, R.: Zukunftssicherheit und Wirtschaftlichkeit beim IT-Einsatz durch eine leistungsfähige Dienstarchitektur – organisatorische und technische Rahmenbedingungen. In: (Lüttich, H.-J.; Rautenstrauch, C., Hrsg.) Verwaltungsinformatik 2000. Verwaltungsinformatik in Theorie, Anwendung und Hochschulausbildung. Mitteldeutscher Verlag, Halle (Saale), 2000, S. 420-424.
- [Er06] Ernst, A.; Lankes, J.; Schweda, C.; Wittenburg, A.: Tool Support for Enterprise Architecture Management – Strengths and Weaknesses. In: The Tenth IEEE International EDOC Conference (EDOC 2006), Hong Kong, 2006.
- [Ge00] Gernert, C.: Architektur als zentrale Aufgabe in heterogenen IT-Systemen. In: (Lüttich, H.-J.; Rautenstrauch, C., Hrsg.) Verwaltungsinformatik 2000. Verwaltungsinformatik in Theorie, Anwendung und Hochschulausbildung. Mitteldeutscher Verlag, Halle (Saale), 2000, S. 47-60.
- [Gr06] Gronau, N.: Wandlungsfähige Informationssystemarchitekturen – Nachhaltigkeit bei organisatorischem Wandel. GITO-Verlag, Berlin, 2006.
- [Ho05] Hoch, D. J.; Klimmer, M.; Leukert, P.: Erfolgreiches IT-Management im öffentlichen Sektor. Gabler Verlag, Wiesbaden, 2005.
- [JW04] Janssen, M.; Wagenaar, R.: Developing Generic Shared Services for e-Government. In: (Bannister, F., Hrsg.) Electronic Journal of e-Government, Academic Conferences Limited, <http://www.ejg.com>; 2004, Volume 2, Issue 1.
- [La05] Lauschke, S.: Softwarekartographie: Analyse und Darstellung der IT-Landschaft eines mittelständischen Unternehmens. 2005. In: <http://www.matthes.in.tum.de/file/Publicationen/2005/Laus05/Laus05.pdf> vom 25.09.2007.
- [Li07] Lietz, G.: Shared Services: E-Government und Economies of Scale. In: (Zechner, A., Hrsg.) Handbuch E-Government. Strategien, Lösungen und Wirtschaftlichkeit. Fraunhofer IRB Verlag, Stuttgart, 2007, S. 269-277.
- [LMW05a] Lankes, J.; Matthes, F.; Wittenburg, A.: Architekturbeschreibung von Anwendungslandschaften: Softwarekartographie und IEEE Std 1471-2000, zitiert nach <http://www.matthes.in.tum.de> vom 20.01.2007. In: Proceedings of the Software Engineering 2005. Essen 2005.
- [LMW05b] Lankes, J.; Matthes, F.; Wittenburg, A.: Softwarekartographie als Beitrag zum Architekturmanagement. In: (Aier, S., Schönherr, M., Hrsg.) Unternehmensarchitekturen und Systemintegration. GITO-Verlag, Berlin 2005, S. 305-333.
- [Ma94] Mattern, K.-H. (Hrsg.): Allgemeine Verwaltungslehre. 4. überarbeitete Auflage, Walhalla Verlag, Berlin, Bonn, Regensburg, 1994.

- [MW04] Matthes, F.; Wittenburg, A.: Softwarekarten zur Visualisierung von Anwendungslandschaften und ihren Aspekten – Eine Bestandsaufnahme. 2004. In: <http://www.matthes.in.tum.de/file/Publikationen/2004/MaWi04a/040326-MaWi-Statusbericht-Softwarekartographie.pdf> vom 25.09.2007.
- [Ni05] Niemann, K. D.: Von der Unternehmensarchitektur zur IT-Governance. Vieweg-Verlag, Wiesbaden, 2005.
- [Sc03] Scheer, A.-W.; Kruppke, H.; Heib, R.: E-Government. Prozessoptimierung in der öffentlichen Verwaltung. Springer, Berlin, Heidelberg, New York, 2003.
- [Sc05] Schönherr, M.: Enterprise Application Integration (EAI) und Middleware, Grundlagen, Architekturen und Auswahlkriterien. In: ERP Management 01 (2005); S. 25-29.
- [TF07] Tschichholz, M.; Fox, O.: Dienste statt Software – Architekturen für prozessorientiertes E-Government. In: (Zechner, A., Hrsg.) Handbuch E-Government. Strategien, Lösungen und Wirtschaftlichkeit. Fraunhofer IRB Verlag, Stuttgart, 2007, S. 133-154.
- [Vo05] Vogel, O., et al: Software-Architektur. Grundlagen – Konzepte – Praxis. München, Elsevier Spektrum Akadem.- Verlag, 2005.
- [Wi06] Wind, M.: IT in der Verwaltung – lange Historie, neue Perspektiven. In: (Wind, M.; Kröger, D., Hrsg.) Handbuch IT in der Verwaltung. Springer, Berlin, Heidelberg, 2006; S. 3-33.

Aufbau und Einsatz der Geschäftsarchitektur bei der AXA Winterthur

Ein minimal invasiver Ansatz

Stephan Aier*, Fiorenzo Maletta**, Christian Riege*, Katja Stucki**, Andrés Frank**

*Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedberg-Strasse 8
CH-9000 St. Gallen
{stephan.aier | christian.riege}@unisg.ch

**AXA Winterthur
General Guisan-Strasse 40
CH-8401 Winterthur
{fiorenzo.maletta | katja.stucki | andrés.frank}@axa-winterthur.ch

Abstract: In diesem Beitrag werden Vorarbeiten zum Aufbau und Einsatz der Geschäftsarchitektur bei der AXA Winterthur beschrieben. Dazu wurde ein minimal invasiver Ansatz gewählt, welcher die Konstruktion eines Geschäftsarchitektur-Metamodells beinhaltet und Aktivitäten im Rahmen eines Vorgehensmodells zur Einbettung der Geschäftsarchitektur in die bestehenden Architekturmanagement-Prozesse beschreibt. Darüber hinaus wird anhand von drei Beispielen demonstriert, wie ein solcher Ansatz zur Geschäftsarchitektur innerhalb der AXA Winterthur genutzt werden kann um die Abstimmung von Geschäfts- und IT-Architekturen zu unterstützen.

1 Einleitung und Motivation

Die AXA Winterthur ist seit dem 1.1.2007 die Schweizer Tochtergesellschaft der AXA. Die AXA Gruppe ist mit 52 Millionen Kunden und einem Umsatz von 79 Mrd. Euro einer der weltweit führenden Versicherungskonzerne. Die AXA Gruppe beschäftigt 150'000 Mitarbeiter und ist in 47 Ländern vertreten. Innerhalb dieser Gruppe repräsentiert die AXA Winterthur die Ländereinheit Schweiz. Tabelle 1 gibt einen Überblick über das Schweizer Geschäft der AXA Winterthur.

AXA Winterthur	
Gründung (als Winterthur Versicherungen)	1875
Anzahl der Privat- und Geschäftskunden	1.8 Mio.
Geschäftsvolumen in CHF (2007)	10.5 Mrd.
Mitarbeitende (per Ende 2007)	4'100
Verwaltetes Vermögen in CHF (per Ende 2007)	62 Mrd.
Marktanteil in der Schweiz (Gesamtwert über alle Sparten)	ca. 20%

Tabelle 1: Unternehmensprofil AXA Winterthur [AW07]

Die Integration der Winterthur Versicherungen in den neuen Rahmenkonzern implizierte auch die Integration in dessen strategische Geschäftsführung. In der AXA Gruppe wurde im Jahr 2007 eine Enterprise Architecture Training Initiative lanciert, welche sicherstellen soll, dass sich die architekturbezogene Integration in die Gruppe auf einer gemeinsamen Basis vollzieht. Die AXA Winterthur startete in demselben Jahr eine Geschäftsarchitektur-Initiative, die im folgenden Beitrag beschrieben wird. Die Initiative ist im Rahmen der folgenden Gesamtziele zu sehen:

- Es soll eine *erhöhte Transparenz* bezogen auf das Unternehmen gewonnen werden. Es gilt, Zusammenhänge, die für die Abwicklung des Geschäfts zentral sind, zu verstehen und zu visualisieren.
- Es soll eine *gemeinsame Sprache* entstehen, welche die relevanten Begriffe als Basis für die unternehmensweite Weiterentwicklung festhält.
- Es gilt, zielgerichtet *Veränderungen* im Geschäftsmodell (i. S. einzelner Anwendungsfälle) zu lokalisieren, um deren Auswirkungen abschätzen zu können. Dies stellt eine wichtige Grundlage für eine koordinierte Weiterentwicklung der Geschäfts- und IT-Landschaften dar.
- Es gilt aufzuzeigen, welche Geschäftsaktivitäten oder Geschäftsprozesse in gleicher Art und Weise ausgeführt werden können. Ziel ist es dabei, eine *Konsolidierung und Standardisierung* (Industrialisierung) der Prozesse unterstützen zu können.

Im vorliegenden Beitrag werden die Vorarbeiten auf dem Weg zur Erreichung dieser Ziele beschrieben. Den zentralen Gegenstand des dabei gewählten Ansatzes bildet das Metamodell der Geschäftsarchitektur als Teil des AXA Winterthur Enterprise Architecture Metamodels. Es ist ein Instrument, um die notwendige Transparenz bezüglich fachlicher Gestaltungsobjekte und ihrer Beziehungen untereinander zu schaffen und eine fachbezogene Weiterentwicklung auch aus Sicht des Business/IT-Alignment zu begleiten. Gleichzeitig unterstützt die geschaffene Transparenz andere Initiativen innerhalb der AXA Winterthur dabei, Synergiepotentiale zu identifizieren sowie ihre strategiekonforme Ausrichtung zu belegen. Dabei wird berücksichtigt, dass die Erstellung der Geschäftsarchitektur ein lebendiger und iterativer Prozess ist und somit die für die Geschäftsarchitektur relevanten Modelle und Prozesse im Wandel begriffen sind. Diese Dynamik wird adressiert, indem der hier vorgestellte minimal invasive Ansatz den Fokus auf ausgewählte Anspruchsgruppen und die Wiederverwendung bzw. auf den Ausbau bereits bestehender Architekturdokumente legt.

Zunächst führt Abschnitt 2 in die begrifflichen Grundlagen zur Unternehmensarchitektur bzw. Geschäftsarchitektur ein. Abschnitt 3 stellt den Ansatz zum Aufbau der Geschäftsarchitektur im Detail vor. Es werden das Metamodell sowie ein exemplarisches Vorgehensmodell für die Gestaltung der Geschäftsarchitektur als die beiden zentralen Bestandteile des Ansatzes präsentiert. Der mögliche Einsatz der Ergebnisse wird in Abschnitt 4 exemplarisch anhand von drei Anwendungsbeispielen demonstriert. Abschnitt 5 gibt einen zusammenfassenden Überblick und stellt weitere geplante Schritte zur Weiterentwicklung der Geschäftsarchitektur bei der AXA Winterthur vor.

2 Grundlagen zum Ansatz der Geschäftsarchitektur

2.1 Verständnis und Positionierung der Geschäftsarchitektur

Architektur wird allgemein definiert als “the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution” [IE00]. Eine Unternehmensarchitektur beschreibt demnach die fundamentale Struktur einer Unternehmung respektive Organisation, zuzüglich dafür notwendiger Gestaltungsprozesse [To07]. In der Regel wird die Unternehmensarchitektur nicht durch eine einzige Architektur beschrieben, sondern durch eine Menge von Teilarchitekturen. Diese können hierarchisch angeordnet werden, bspw. als eine Menge von Subsystemen, durch die Aufteilung in Geschäftsbereiche [Ba00; Le04] oder die Anordnung in Form von konzeptionellen Ebenen [To07; WF07; Za87]. Die Unternehmensarchitektur vernetzt bestehende Modelle, wie etwa Produktmodelle, Geschäftsprozessmodelle, Domänenmodelle, Organigramme, Arbeitsplatzbeschreibungen oder Funktionsbäume zu einem konsistenten, umfassenden Abbild der Organisation. Hiervon ausgehend besitzt bereits jedes Unternehmen eine Unternehmensarchitektur, allerdings wird sie nicht in jedem Unternehmen genutzt [Ni05].

Die Unternehmensarchitektur als zweckorientiertes Abbild der Organisation fungiert

- als Kommunikationsbasis und Dokumentation des Ist-Zustandes der Unternehmensstrukturen und Prozesse,
- als Unterstützung für die Erstellung des Soll-Zustandes der Unternehmensstrukturen und Prozesse und
- als Unterstützung von Transformationsprojekten vom Ist- zum Soll-Zustand [FW07; Pa06].

In Abgrenzung zum vorgestellten Unternehmensarchitekturverständnis positioniert sich die Geschäftsarchitektur als diejenige Teilarchitektur, welche die fachlichen Gestaltungsobjekte einer Organisation implementierungsneutral abbildet [To07]. Sie beinhaltet dazu Grundstrukturen bzgl. der Geschäftsziele, Geschäftsprozesse und Organisationsstrukturen sowie dafür notwendige Ressourcen [Ke07; VB06; We07; WM04]. Des Weiteren stellt sie die Beziehungen dieser Elemente untereinander dar. Als besonders kritisch wird oftmals die Schnittstelle zwischen fachlichen Strukturen und Informationssystemen angeführt. Die Geschäftsarchitektur berücksichtigt diese Schnittstelle indem Zusammenhänge bspw. zwischen Geschäftsprozessen und ihrer Unterstützung durch bestimmte Applikationen abgebildet sind. Die Unternehmensarchitektur, und somit auch die Geschäftsarchitektur, fungiert als ein zentrales Koordinationsinstrument für das Business/IT-Alignment [FW07].

2.2 Gestaltung des Geschäftsarchitektur-Modells

Für die (Aus-)Gestaltung der Geschäftsarchitektur existiert eine Vielzahl von Ansätzen. Standardisierte Facharchitekturmodelle wie IAA [HR06] oder VAA [Ge01] helfen dabei, das eigene Verständnis hinsichtlich abgebildeter Gestaltungsobjekte und Zusammenhänge auf Vollständigkeit und Anschlussfähigkeit innerhalb der Versicherungsbranche zu schärfen. Sie können jedoch selten unverändert auf die Gegebenheiten einer spezifischen Organisation übertragen werden. Ebenso liefern Frameworks zur Architekturgestaltung wie das Zachmann-Framework, das US Federal Government Enterprise Architecture Framework (FEAF) oder The Open Group Architecture Framework (TOGAF) wertvolle Hinweise für ein zielgerichtetes und ganzheitliches Vorgehen bei der Erstellung und Pflege von Architekturen [FE07; To07; Za87]. Die konkrete Ausgestaltung einzelner Prozessschritte muss sich jedoch an den spezifischen Gegebenheiten der Organisation und den mit der Geschäftsarchitektur verfolgten Zielstellungen orientieren. Für die hier vorgestellte Initiative zur Geschäftsarchitektur ergeben sich aus diesen Überlegungen drei zentrale Fragestellungen:

- Welche Gestaltungsobjekte der Geschäftsarchitektur sind zu berücksichtigen, um die in Abschnitt 1 aufgeführten Ziele zu erreichen, Zusammenhänge für die relevanten Anspruchsgruppen abzubilden sowie dem Wunsch nach Anpassbarkeit nachzukommen?

- Welches leichtgewichtige und iterative Vorgehen zur Erstellung und Pflege der Geschäftsarchitektur ist zu wählen, um alle relevanten Anspruchsgruppen zu erreichen?
- Welche Gestaltungsprinzipien sind für Akteure leitend, die implizit oder explizit, in formalem oder informellem Rahmen, die Architektur mit Leben füllen? Beispiele für Akteure sind Prozessmanager, Linienmanager, Business Engineers, Business Analysten und Solution Architekten.

Ein Kernbestandteil der Unternehmensarchitektur und damit auch der Geschäftsarchitektur ist ein Metamodell, welches die Quintessenz des Gestaltungsgegenstandes „Unternehmen“ abbildet [Oe07]. Im Metamodell der Geschäftsarchitektur werden demnach die aggregierten Gestaltungsobjekte der relevanten Teilarchitekturmodelle definiert und zueinander in Beziehung gesetzt werden [WF07]. Mit Realdaten befüllt, ist das Geschäftsarchitekturmodell ein wichtiges Instrument, um die in Abschnitt 1 aufgelisteten Zielstellungen zu adressieren. Es kann u. a. für folgende Aufgaben herangezogen werden [De07; Ki07]:

- Werden Entscheidungen auf strategischer, taktischer oder operativer Ebene vorbereitet so unterstützt die Geschäftsarchitektur dabei, dass diese konform mit den grundsätzlichen Unternehmenszielen sind.
- Die Geschäftsarchitektur eignet sich insbesondere für Dokumentations- und Schulungszwecke, um z. B. das Verständnis bei Strategie- und Implementierungsdiskussionen zu vereinheitlichen.
- Die Geschäftsarchitektur hilft, Ansatzpunkte für Prozessveränderungen und -automatisierungen zu identifizieren, sowie Wechselwirkungen zwischen dem Geschäftsmodell und der Informationstechnologie transparent darzustellen.

Die Konstruktion eines Geschäftsarchitektur-Metamodells darf nicht losgelöst von der Erst- und Weiterentwicklung der Geschäftsarchitektur erfolgen. Vielmehr muss sie als zentraler Bestandteil in einer Methode zur Erstellung und Pflege der Geschäftsarchitektur eingebettet sein. Basierend auf den Konzepten zum Method Engineering [Br06; KW92; LV07] beschreibt eine Methode dafür systematisch ein Vorgehen, wie von einer problembehafteten Ausgangslage zu einem problemadäquaten Ergebnis zu gelangen ist [Be01]. Um dem Anspruch an einen leichtgewichtigen Ansatz gerecht zu werden, verknüpft ein Vorgehensmodell die Metamodellkonstruktion mit der Berücksichtigung von Anspruchsgruppen und bettet alles in ein bestehendes Architekturmanagement ein. Ebenfalls beinhaltet es die Kommunikation der Ergebnisse in Form von Anwendungsbeispielen der Geschäftsarchitektur.

3 Ansatz der minimal invasiven Geschäftsarchitektur

3.1 Verständnis und Einbettung

Ein Ansatz zum Aufbau und Einsatz der Geschäftsarchitektur bei der AXA Winterthur muss sich sinngemäß zum Einen mit der Konstruktion eines entsprechenden Metamodells zur Abbildung der relevanten fachlichen Gestaltungsobjekte auseinandersetzen. Zum Anderen ist die Definition eines geeigneten Vorgehensmodells zur Einführung, Verankerung und kontinuierlichen Bewirtschaftung der Geschäftsarchitektur innerhalb der Organisation notwendig. Grundlegend für das Metamodell und das Vorgehensmodell sind Gestaltungsprinzipien zu definieren, welche durch ihren normativen Charakter für die effektive Gestaltung des Unternehmens leitend sein sollen. In allen drei Punkten wird Wert darauf gelegt, dass ihre Bearbeitung möglichst keine tiefgreifende Umgestaltung innerhalb der Organisation nach sich zieht, im Sinne eines minimal invasiven Ansatzes. Minimal invasiv bedeutet, dass das Konzept der Geschäftsarchitektur ohne organisatorische Restrukturierungen und Aufgabenveränderungen etabliert werden kann. So wird gewährleistet, dass das Konzept anschlussfähig zu existierenden Architekturmanagement-Prozessen ist. Darüber hinaus sollen bestehende Architekturergebnisse und -dokumente beim Aufbau der Geschäftsarchitektur herangezogen werden. Ebenfalls wichtig ist, dass der Ansatz zum Aufbau der Geschäftsarchitektur iterativen Charakter besitzt, um zukünftige Änderungen im Geschäftsmodell oder neue Anforderungen seitens relevanter Anspruchsgruppen aufnehmen zu können.

Angewandt auf die Geschäftsarchitektur, fördert ein minimal invasiver Ansatz somit die Verbreitung und die Akzeptanz derselben innerhalb der Organisation. Erfahrungen und weitergehende Ansätze finden sich u. a. im Thema Organizational Awareness [MZ07]. Als zentrales Element verkörpern die Gestaltungsprinzipien Leitsätze, nach denen sich die Ausgestaltung der Geschäftsarchitektur im Sinne der ihr zugeordneten Modelle und ihrer Prozesse zur Bewirtschaftung richtet. Innerhalb der AXA Winterthur sind diese allgemeinen Gestaltungsprinzipien definiert und bilden den Rahmen für den vorgestellten Ansatz. Abbildung 1 stellt die drei zentralen Elemente für die Gestaltung der Geschäftsarchitektur bei der AXA Winterthur dar.

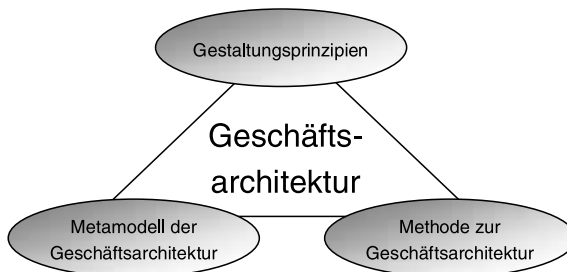


Abbildung 1: Zentrale Elemente für den Ansatz der Geschäftsarchitektur

Nachfolgend wird zum einen das konstruierte Geschäftsarchitektur-Metamodell anhand der wichtigsten Gestaltungsobjekte beschrieben. Zum anderen wird das Vorgehen, als das zentrale Element einer Methode, zum Aufbau der Geschäftsarchitektur anhand der wichtigsten Aktivitäten dargestellt.

3.2 Metamodell der Geschäftsarchitektur

Für die Erstellung des Geschäftsarchitektur-Metamodells wurde ein pragmatisches Vorgehen gewählt, indem als Grundlage das Unternehmensarchitektur-Metamodell der AXA Winterthur herangezogen wurde. Der Zweck des Unternehmensarchitektur-Metamodells besteht in der Erlangung einer zentralen, standardisierten und zusammenhängenden Sichtweise über das Unternehmen als Ganzes. Das Metamodell beschreibt aus Sicht der Geschäftsarchitektur relevante Elemente, Beziehungen, Attribute sowie deren Semantik und Qualitätsmerkmale. Es galt zunächst, die geschäftsarchitekturrelevanten Entitätstypen und Beziehungen zu identifizieren. Dazu sind die vorgängig aufgenommenen Gestaltungsobjekte und Lieferobjekte in entsprechende Elemente für das Geschäftsarchitektur-Metamodell zu überführen, vgl. dazu Abschnitt 3.3. Die methodische Grundlage und Beispieltransformationen bietet in diesem Zusammenhang das Viewpoint based Metamodel Engineering [KW07]. Die nachfolgende Auflistung gibt ausgewählte Entitätstypen des Geschäftsarchitektur-Metamodells der AXA Winterthur wieder.

- *Value Chain*. Die Value Chain kategorisiert und konsolidiert die strategisch wichtigen Aktivitäten des Unternehmens. Es wird unterschieden in Value-adding-, Management- und Support-Aktivitäten. Die Value Chain stellt einen Ansatz dar, der systematisch auf die Entwicklung von Tätigkeiten fokussiert, die ausschlaggebend sind für die Erlangung von Wettbewerbsvorteilen.
- *Business Capabilities*. Eine Business Capability repräsentiert die Fähigkeit oder Fertigkeit eines Unternehmens, Wert steigernd zu agieren. Eine Business Capability kapselt dazu Personen, Geschäftsprozesse und Technologien in Verbindung mit einer entsprechenden Geschäftsfunktion.
- *Product*. Ein Produkt, definiert als rechtliche bzw. vertragliche Einheit, ist das was die AXA Winterthur auf dem Markt anbietet. Es ist das Leistungsversprechen, das die AXA Winterthur dem Kunden abgibt (struktureller Aspekt). Gleichzeitig ist das Produkt eine Leistungserbringung (Verhaltensaspekt) [Fa00].
- *Business Process*. Ein Geschäftsprozess ist eine Abfolge von Geschäftsaktivitäten, die über mehrere Sparten verteilt sein können und in einer vorgegebenen Ablauffolge zu erledigen sind. Im Rahmen eines Geschäftsprozesses werden mit unternehmensinternen oder -externen Kunden (Prozess-)Leistungen ausgetauscht. Geschäftsprozesse sind unmittelbar auf die Stiftung von Kundennutzen ausgerichtet. Sie werden u. a. über Prozesskennzahlen gesteuert und weiterentwickelt [Oe95].

Eine Auswahl an Kernentitäten der Geschäftsarchitektur zeigt Abbildung 2. Dabei soll der grundsätzliche Zusammenhang zwischen den Entitäten dargestellt werden. Insbesondere wird aufgezeigt, in welcher Verbindung die Entitäten der Geschäftsarchitektur zu den Entitäten der IS-Landkarte stehen.

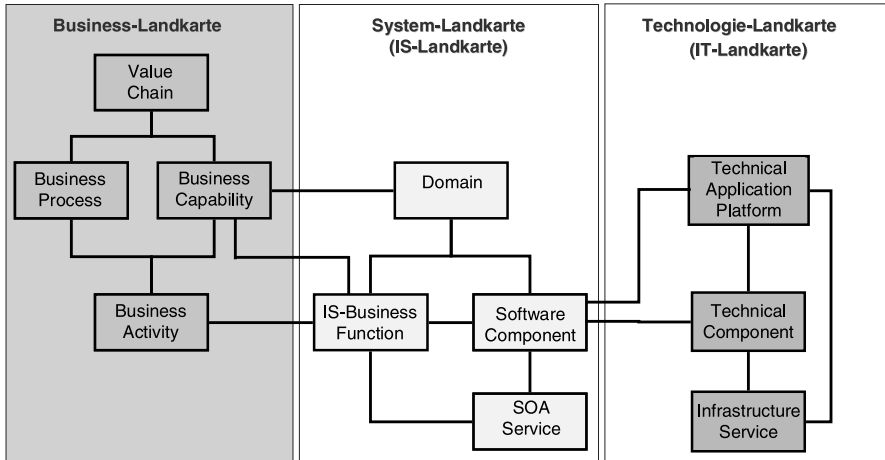


Abbildung 2: High-level Sicht der Geschäftsarchitektur (Business Landkarte) und ihre Verbindung zur IS-Landkarte

Das Geschäftsarchitektur-Metamodell der AXA Winterthur unterscheidet verschiedene Abstraktionsebenen sowie einzelne thematische Sichten auf die Geschäftsarchitektur. Dies erfolgt aus dem Bedürfnis heraus, jeder Anspruchsgruppe der Architektur diejenigen Sichten zur Verfügung zu stellen, die ihre Anliegen illustrieren [IE00; La05]. Dem minimal invasiven Ansatz folgend wird dadurch außerdem gewährleistet, dass Änderungen im Modell nicht stets modifizierte Darstellungen für alle anderen Anspruchsgruppen zur Folge haben. So wurden für die Geschäftsarchitektur der AXA Winterthur u. a. folgende Metamodell-Sichten erarbeitet:

- *Strategiesicht.* Die Strategiesicht zeigt, wie das Unternehmen strategisch gesteuert wird. Dafür sind Business Imperatives als diejenigen Treiber dargestellt, die hinter der konkreten Ausprägung einer Geschäftsstrategie stehen. Die Geschäftsstrategie gibt Anhaltspunkte dafür, welche Business Capabilities entwickelt werden sollen. Deren Entwicklung geschieht dann in Form von strategischen Initiativen.
- *Geschäftsprozesssicht.* Ziel der Sicht ist es, ein gemeinsames Verständnis der Anspruchsgruppen über die Zusammenhänge von Geschäftsprozessen und Business Capabilities, Aufbau- und Ablauforganisation zu schaffen. Ausserdem wird die Bedeutung betrieblicher Kennzahlen deutlich. Die Sicht dient u. a. zur Verbesserung von Templates für Prozessbeschreibungen, zum Ableiten von Checklisten für die Beschreibungen von Prozessen, Aktivitäten und Rollen oder zur Definition von Kennzahlen.

- *Domänensicht.* Die Domänensicht ist die Grundlage für die Erstellung der Domänen- und Applikationslandkarte der AXA Winterthur. Die Domänensicht verdeutlicht den Übergang zwischen Business und IS-Landschaft. Die Sicht zeigt, wie IT-unterstützte Geschäftsfunktionen und Entitäten der Informationslandschaft gruppiert werden. Das Gruppierungskriterium entspricht dabei dem Business Capability-Modell der AXA Gruppe.

Die unterschiedlichen Abstraktionsebenen betonen sowohl eine funktionsorientierte Sicht, wie auch eine prozessorientierte Sicht auf die Geschäftsarchitektur. Eine herausgehobene Position besitzen die Business Capabilities, als Element der funktionsorientierten Sicht. Für die prozessorientierte Sicht sind dies äquivalent die Geschäftsprozesse. Während Geschäftsprozesse eine Sequenz der Aktivitäten zeigen, ist die funktionsorientierte Sicht eine Sammlung von für die Geschäftsabwicklung wichtigen Fähigkeiten. Diese Sammlung ist im Vergleich zu Prozessen beständiger. So werden u. a. stabile Funktionsbausteine in Form von Capabilities auf die benötigten informationstechnischen Ressourcen abgebildet. Aus einer prozessorientierten Perspektive sind Entitäten wie Rolle und Organisation von Bedeutung, welche wiederum aus einer funktionalen Perspektive weniger bedeutsam ist.

3.3 Vorgehensmodell zur Erstellung und Pflege der Geschäftsarchitektur

Das zweite zentrale Ergebnis des Ansatzes ist das Vorgehensmodell. Es ordnet die Aktivitäten zur Erstellung des Metamodells der Geschäftsarchitektur und beschreibt die Aktivitäten, welche für die Pflege und Nutzung der Geschäftsarchitektur notwendig sind. Als Grundlage wurde das Vorgehen zur Architekturdefinition gemäß TOGAF gewählt [To07]. Die Orientierung an den Phasen der Architecture Development Method stellt sicher, dass die Aktivitäten zur Erstellung und Pflege der Geschäftsarchitektur der AXA Winterthur dem gewählten Standard der AXA Gruppe entsprechen und ihre Abfolge im Gesamtprozess nachvollziehbar ist. Es fördert zugleich die Anschlussfähigkeit der gewonnenen Ergebnisse in den einzelnen Aktivitäten. Das resultierende Vorgehensmodell, dargestellt in Abbildung 3, zeigt die einzelnen Aktivitäten.

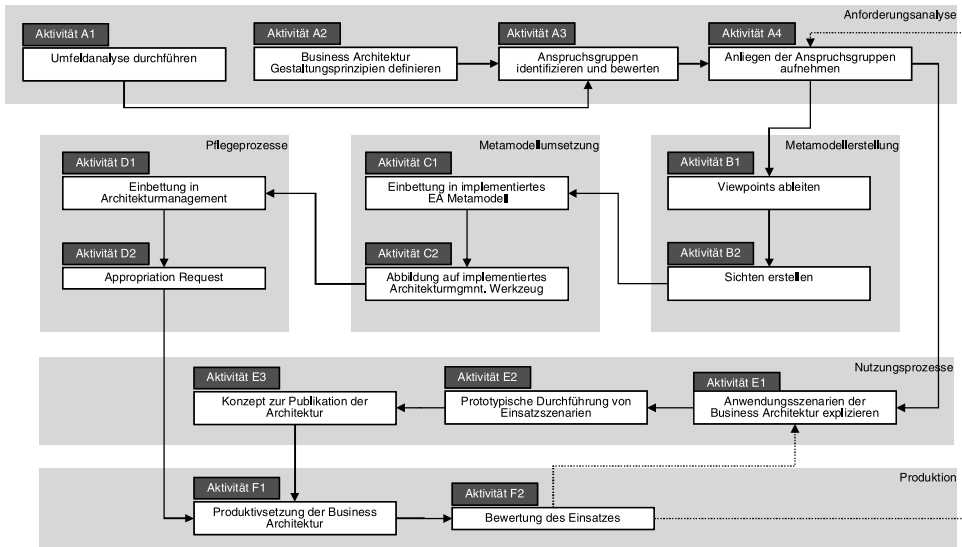


Abbildung 3: Vorgehensmodell zur Erstellung und Pflege der Geschäftsarchitektur

Das Vorgehensmodell ist in sechs verschiedene Phasen gegliedert. Innerhalb einer jeden Phase sind verschiedene Aktivitäten eingeordnet, welche nachfolgend beschrieben sind.

- *Phase der Anforderungsanalyse.* Es gilt, eine Umfeldanalyse durchzuführen. Hier werden vergleichbare Initiativen und Standardisierungsvorschläge wie IAA und VAA analysiert, und es werden ggf. Erfolgsfaktoren der Geschäftsarchitektur abgeleitet. Des Weiteren sind die Anspruchsgruppen der Geschäftsarchitektur zu identifizieren und mit Hilfe einer Zielkonfliktmatrix zu analysieren. Im Sinne des minimal invasiven Ansatzes kann es sinnvoll sein, Anspruchsgruppen zeitlich versetzt an das Thema Geschäftsarchitektur heranzuführen [Sc05]. Es sind ebenso die Informationsbedarfe der Anspruchsgruppen aufzunehmen, die mit Hilfe der Geschäftsarchitektur adressiert werden sollen.
- *Phase der Metamodellerstellung.* Für die Abbildung der Anliegen der Anspruchsgruppen im Geschäftsarchitektur-Metamodell müssen die entsprechenden Entitätstypen und Beziehungen identifiziert werden. Eine einzelne, übergreifende Repräsentation der Geschäftsarchitektur ist oftmals zu komplex, um alle Entitäten und Beziehungen auf hohem Detaillevel zu verstehen und zu kommunizieren. Daher werden verschiedene Sichten benötigt, die die gewünschten Anliegen der Anspruchsgruppen visualisieren, vgl. Abschnitt 4.1.
- *Phase der Metamodellumsetzung.* Der Ausschnitt des Geschäftsarchitektur-Metamodells muss in ein existierendes Unternehmensarchitektur-Metamodell integriert werden. Hierfür wird der Pflegeprozess für das Unternehmensarchitektur-Metamodell um hinzugefügte Sichten der Geschäftsarchitektur erweitert. Weiterhin ist es notwendig, das bestehende Glossar zum Metamodell auf Konsistenz bzgl. der erweiterten Modellbestandteile zu überprüfen. Gegebenenfalls muss das Begriffssystem mit neuen bzw. modifizierten Entitätstypenbeschreibungen ergänzt werden.

- *Phase der Pflegeprozesse.* Diese Aktivität definiert die Anknüpfungspunkte der Geschäftsarchitektur aus organisatorischer und technischer Sicht als Teil des übergeordneten Architekturmanagements [De06; Ha05]. Im Sinne des minimal invasiven Ansatzes werden bestehende Prozesse des Architekturmanagements, Rollen- und Rechtekonzepte nur dort angepasst, wo Handlungsbedarf aus Sicht der Geschäftsarchitektur und ihrer Anspruchsgruppen besteht und diese Anpassung möglich ist. Dabei ist ein behutsames, u. U. schrittweises Vorgehen zweckmäßig, um das Potential der Geschäftsarchitektur punktuell sichtbar zu machen und dabei anschlussfähig zu bleiben.
- *Phase der Nutzungsprozesse.* Ein Ziel dieser Aktivität ist es, Anwendungsbereiche zu identifizieren, in denen die Arbeit mit der Geschäftsarchitektur relevant ist oder in denen Teile der Geschäftsarchitektur konkret ausgearbeitet oder weiterentwickelt werden. Dafür werden bestehende Projekte bzw. Initiativen innerhalb der Unternehmung identifiziert, in deren Kontext die Geschäftsarchitektur mit einiger Wahrscheinlichkeit etwas bewirken kann. Die Menge an fachlichen Anwendungsszenarien der Geschäftsarchitektur ergibt sich dabei aus den zuvor erhobenen Anliegen der Anspruchsgruppen, siehe dazu Abschnitt 4. Es genügt, zunächst wenige Anwendungsszenarien zu identifizieren. Wichtig ist, dass sie dem Gedanken von Geschäftsarchitektur bereits positiv gegenüberstehen [To07]. Ebenfalls gilt es, das Feedback der Anspruchsgruppen zu erfassen und das Ergebnis anhand der spezifizierten Anliegen zu messen.
- *Phase der Produktion.* Im Rahmen der Produktivsetzung der Geschäftsarchitektur erfolgt zum einen die initiale Befüllung des Repositories mit Produktivdaten, zum anderen werden die identifizierten Sichten in einem entsprechenden Repository-Ausschnitt erstellt. Es gilt die Anwendung der Geschäftsarchitektur (Metamodell, Prozesse) zu bewerten, um ggf. iterativ Ausbau bzw. Anpassungen der Geschäftsarchitektur vorzunehmen. Dazu können Prinzipien und Metriken zur Beurteilung von Architekturen aufgestellt werden [Sc05]. Die Schlussfolgerungen dieser Aktivität fließen entweder in die Anwendung der Geschäftsarchitektur ein oder sind Grundlage für eine Anpassung des Metamodells.

Das Konzept der Geschäftsarchitekturen, soll es dauerhaft verankert werden, kann nur schrittweise in die Organisation getragen werden. Dieser Erkenntnis wird das vorgestellte Vorgehensmodell gerecht. Ebenfalls unterstützt das vorgeschlagene Vorgehen die kontinuierliche Weiterentwicklung der Geschäftsarchitektur, indem es einen Abgleich mit definierten Gestaltungsprinzipien vorsieht. Im folgenden Abschnitt wird das Nutzenpotential der Geschäftsarchitektur exemplarisch anhand von drei Einsatzszenarien in der AXA Winterthur vorgestellt.

4 Anwendungsbeispiele für den Einsatz der Geschäftsarchitektur

4.1 Geschäftsarchitektur zur Unterstützung der strategischen Unternehmensführung

Die strategische Unternehmensführung befasst sich mit der langfristigen und grundsätzlichen Ausrichtung des Gesamtversicherungsunternehmens. Sie koordiniert dazu die großen Unternehmensbereiche entlang der Wertschöpfungskette [Fa00]. Eine Grundlage für die Strategiearbeit der AXA Winterthur bilden acht strategische Initiativen, welche die drei übergreifenden Stossrichtungen adressieren. Diese drei Stossrichtungen sind derzeit: *Stärkung des bestehenden Geschäfts*, *Erschließung neuer Geschäftsfelder* und *Berücksichtigung von Mitarbeiter-Engagement und Kundenfokus*. Die Herausforderung in der strategischen Unternehmensführung berührt die Arbeit zur Geschäftsarchitektur u. a. in Form folgender Fragestellungen:

- Welche Geschäftsfunktionen sollen aufgrund veränderter Rahmenbedingungen in einem bestimmten Markt hinsichtlich Kosten, Qualität und Verarbeitungsgeschwindigkeit angepasst werden?
- Ist es möglich, bei gleich bleibender Qualität die Anzahl der Kunden für eine Marktleistung zu erhöhen?
- Kann eine bestimmte Marktleistung auch für andere Kundensegmente angeboten werden?
- Kann eine bestimmte Marktleistung in Zukunft als Leistungsbündel angeboten werden?

Unternehmensarchitektur als Container der Geschäftsarchitektur stellt zu diesem Zweck eine konzeptionelle Sichtweise auf strategische Initiativen dar. Dazu ist es notwendig, dass die strategischen Ziele und ein Plan zur deren Erreichung definiert sind. An dieser Stelle kann die Geschäftsarchitektur als ein Analyseinstrument für die strategische Unternehmensführung genutzt werden [De07]. Mit ihrer Hilfe ist es u. a. möglich, die Geschäftsstrategie in Form formulierter Initiativen auf die Business Capabilities abzubilden. Abbildung 4 stellt hier exemplarisch eine solche Zuordnung von Initiativen auf ausgewählte Elemente des Business Capability-Modells dar.



Abbildung 4: Einfluss von Geschäftsinitiativen auf Business Capabilities

Darauf aufbauend kann das Projektportfolio bezüglich seiner Strategiekonformität beurteilt werden, indem Abhängigkeiten zwischen Projekten und die durch einzelne Projekte adressierten strategischen Initiativen erkannt werden. Es ist vorstellbar, dass zwei Initiativen dieselbe Business Capability, z. B. *Manage sales partners and alliances* weiterentwickeln. Es soll ebenso identifiziert werden, ob zwei Initiativen die Kapazität derselben IT-Abteilung beanspruchen, weil beide Modifikationen an denselben Applikationen bedingen. Die Folge wäre, dass die Initiativen zu priorisieren sind. Umgekehrt ist es denkbar, dass zwei Initiativen sich in ihrer Absicht überschneiden oder große Übereinstimmungen aufweisen, jedoch verschiedene Teile der Applikationslandschaft adressieren. Es sind Überlegungen notwendig, inwiefern Redundanz vermieden werden kann, möglicherweise bevor die Initiativen gestartet werden. Im Ergebnis ist es mit Unterstützung der Geschäftsarchitektur möglich, das Projektportfolio entsprechend der Geschäftsstrategie zu beurteilen und im selben Zug das Applikationsportfolio aktiv zu bewirtschaften.

4.2 Geschäftsarchitektur und Produktentwicklung

Das Versicherungsprodukt, im Sinne einer Marktleistung, wird durch Informationen abgebildet. Aus diesem Grund weist die Versicherungsschutzproduktion eine starke informationstechnische Komponente auf [Fa00]. Die Produktentwicklung ist ein zentrales Thema in der Versicherungswirtschaft; ausgehend von früheren Segmentierungsstrategien, wie der Trennung nach Sparten, wurden zahlreiche Legacy-Systeme für die Produktdefinition und Produktverwaltung aufgebaut [Ro04]. Eine aktuelle Herausforderung ist es demnach, in Systemen den Austausch von Informationen über Unternehmensbereiche und Organisationseinheiten zu ermöglichen. Dazu können u. a. folgende für die Geschäftsarchitektur relevanten Prinzipien formuliert werden:

- Die IT-Beteiligung soll sich auf ein Minimum beschränken und die Fachseite in die Lage versetzen, Produkte zu entwickeln und zu verwalten.
- Die Produktmerkmale sollen separiert von Vertragsverwaltungssystemen definiert und vorgehalten werden.
- Die Aktivitäten im Produktmanagement sollen verstärkt zentral abgewickelt werden, um den Fokus auf den Kundenwunsch zu legen.

Das Funktionscluster Produktentwicklung umfasst die Produktdefinition, Produktpflege und Beitragsabrechnung [Ro04]. Die Geschäftsarchitektur kann zunächst einmal Transparenz schaffen, indem sie die verschiedenen Sichten rund um das Gestaltungsobjekt Produkt, wie Kundenproduktsicht, vertriebsorientierte, tarifliche und Produktentwicklersicht miteinander in Beziehung setzt. Es wird dargestellt, welche Elemente einen Einfluss auf das Versicherungsprodukt haben, d. h. z. B. welche Attribute ein bestimmtes Produkt besitzt oder welche Rollen am Produktentwicklungsprozess beteiligt sind. Es wird ebenso deutlich, welche Prozesse und Applikationen im Rahmen der Produktentwicklung involviert sind und welche Abhängigkeiten zwischen ihnen bestehen. Bezogen auf die Dimensionen der Gestaltung von Versicherungsprodukten können u. a. folgende Fragestellungen von Interesse sein:

- Wie ist das Produkt positioniert, bezogen auf das Produktportfolio?
- Über welche Absatzkanäle wird das Produkt abgesetzt und welches Markt- bzw. Kundensegment kann angesprochen werden?
- Welche ggf. versicherungsfremden Zusatzleistungen sind Bestandteil des Produktes und wie ist die Schnittstelle zum Geschäftspartnerprozess?

Steht ein flexibles Produktmodell zur Verfügung, so kann das dazugehörige Lebenszykluskonzept von Produkten dem Business Capability-Modell gegenübergestellt werden, um funktionale Anforderungen aus Produktsicht mit den zentralen, definierten Fähigkeiten der AXA Winterthur abzugleichen. Es ist denkbar, Applikationen dahingehend zu identifizieren, an welchen Punkten sie im Produktlebenszyklus genutzt werden. Somit wird gewährleistet, dass die erforderlichen Informationen für die Produktentwicklung auch zur Verfügung stehen, kritische Prozesse bekannt sind, sowie verwendete Daten nicht redundant vorgehalten werden.

4.3 Werkzeugunterstütztes Management der Geschäftsarchitektur

In den geschäftlichen Abläufen der AXA Winterthur wird eine Vielzahl von Informationen produziert oder bearbeitet. Verstreut über das ganze Unternehmen pflegen einzelne Organisationseinheiten und Projekte oft individuelle Vorlagen und Arbeitsergebnisse wie z. B. Briefvorlagen, Excel-Listen oder auch Dokumentenablagen. Hinter jeder Vorlage und hinter jedem Ergebnis liegt ein Verständnis darüber, wie solche Vorlagen oder Arbeitsergebnisse aufgebaut sein sollen. Bestenfalls sind solche Ad-hoc-Verständnisse dokumentiert, in Form von Glossaren, Richtlinien oder Templates. Oftmals sind sie jedoch nur implizit vorhanden. Folgende Phänomene können somit häufig beobachtet werden:

- Eine Vielfalt ähnlicher Ergebnisse mit Bezug zur Geschäftsarchitektur und unterschiedlicher Ausprägung wird erzeugt und gepflegt.
- Diskussionen um dieselben Themen wiederholen sich und Zusammenhänge zwischen voneinander abhängigen oder ähnlichen Ergebnissen werden nicht erkannt.

Um dem entgegenzuwirken, werden Ergebnisse als Output von Prozessen betrachtet. So wird ersichtlich, welche Ergebnisse in welchen Aktivitäten produziert und wo sie weiterverwendet werden. Dies ist eine Voraussetzung für die Erstellung von gemeinsam verwendeten Templates und Dokumentstrukturen. Das Metamodell der Unternehmensarchitektur, im speziellen der Geschäftsarchitektur, visualisiert diese Überlegungen und dient als Vorlage zur Ausgestaltung von Repositories.

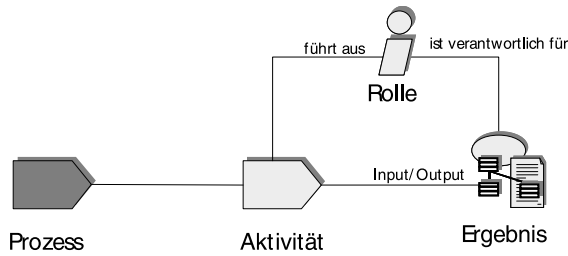


Abb. 5: Vereinfachtes Metamodell für Prozesse

Für den gezeigten Ausschnitt der Geschäftsarchitektur bspw. wäre dies das Prozessmodellierungswerkzeug. Es muss gestatten, Sequenzen von Aktivitäten darzustellen, Aktivitäten Rollen zuzuordnen, Input und Output von Aktivitäten als Ergebnisse zu definieren und gesamthaft zu Prozessen zu bündeln. Optimal ist, wenn die Struktur eines solchen Werkzeugs, d. h. das zugrunde gelegte Metamodell sich am Geschäftsarchitektur-Metamodell orientiert. Ein werkzeugunterstütztes Management der Geschäftsarchitektur zielt auf die Speicherung von Metainformationen über Elemente wie Business Capabilities oder Informationssysteme in einem Repository ab [TL04].

Wie in Abschnitt 4.1 beschrieben, werden die strategischen Initiativen auf Business Capabilities abgebildet und es wird untersucht, welche Business Capabilities durch welche Applikationen unterstützt werden. Applikationen sowie die Systemsoftware und Hardware, welche die technische Infrastruktur dieser Software darstellen, sind Beispiele von Architektur-Assets der AXA Winterthur. Diese Assets benötigen ein Lifecycle Management, welches optimalerweise werkzeugunterstützt betrieben wird, um Zusammenhänge und Abhängigkeiten automatisiert aufzudecken. Zusätzlich können u. a. automatisiert Kennzahlen generiert werden, zum Zweck einer optimalen Bereitstellung von IT für die Bedürfnisse der Fachbereiche. Dies bedingt jedoch, dass die Qualität der erfassten Daten gewährleistet wird und Ressourcen bereitstehen, um diese Daten zu pflegen und aktuell zu halten. Das Geschäftsarchitektur-Metamodell unterstützt zusätzlich dabei, die Ausdehnung von Pflege- und Nutzungsprozessen in andere Landschaftsgebiete bewusst zu steuern. Eine solche Ausdehnung wurde exemplarisch versucht, indem in Ergänzung zu den bereits gepflegten IT- und IS-Daten neu auch Daten der Business-Landschaft, vgl. Abschnitt 3, erfasst und mit den Applikations-Daten verknüpft wurden. Die Verknüpfung zur IS-Landschaft geschah exemplarisch anhand eines größeren Umbau-Projektes im Bereich Produkt und Vertrag.

5 Zusammenfassung und weitere geplante Schritte

Der Praxisbeitrag stellt einen Ansatz zum Aufbau und für den Einsatz der Geschäftsarchitektur bei der AXA Winterthur vor. Ziel dieses als minimal invasiv charakterisierten Ansatzes ist es dabei

- das Verständnis des Geschäftsmodells bzw. Unternehmensmodells der AXA Winterthur zu schärfen,

- eine gemeinsame Sprache und relevante Begriffe als Instrument für die unternehmensweite, als auch Ländereinheiten übergreifende Weiterentwicklung zu etablieren,
- sowie mögliche Änderungen im Geschäftsmodell frühzeitig lokalisieren und deren Auswirkungen abzuschätzen.

Als zentrale Bestandteile des Ansatzes wurde dabei ein Geschäftsarchitektur-Metamodell konstruiert sowie ein Vorgehensmodell zu Einführung der Geschäftsarchitektur definiert. Das Geschäftsarchitektur-Metamodell erweitert das bestehende Unternehmensarchitektur-Metamodell um geschäftsarchitekturrelevante Gestaltungsobjekte. Das Vorgehensmodell zur Erstellung und Pflege der Geschäftsarchitektur verknüpft einzelne Aktivitäten, wie die Definition von Anspruchsgruppen der Geschäftsarchitektur, die Konstruktion des Metamodells inklusive benötigter Detailsichten sowie die Einbettung in ein übergeordnetes Architekturmanagement. Zusätzlich schließt das Vorgehensmodell die exemplarische Anwendung der Geschäftsarchitektur ein. Die Anwendungsbeispiele unterstreichen, dass die Geschäftsarchitektur aus Perspektive der strategischen Unternehmensführung, im Rahmen der Produktentwicklung als auch zur Automatisierung von Auswertungen und Analysen einen Beitrag leisten kann. Erfolgsfaktoren waren dabei u. a. die frühzeitige und kontinuierliche Einbindung der definierten Anspruchsgruppen im Rahmen der Metamodellkonstruktion als auch bei der Auswahl und Durchführung von Anwendungsbeispielen. Ebenso soll der Fokus bei der Implementierung von ersten Anwendungen der Geschäftsarchitektur auf ihre einfache und wirksame Durchführbarkeit gelegt werden. Im Verständnis des hier vorgestellten minimal invasiven Ansatzes ist ein real vorzeigbares, plausibles Resultat im Sinne des Geschäftsarchitektur-Marketings als wichtig einzustufen. Weitere geplante Ausbauschritte der Geschäftsarchitektur bei der AXA Winterthur liegen u. a. in der Nutzung im Rahmen eines Ausbildungs- und Anforderungsprofils für Business Analysten.

6 Literaturverzeichnis

- [AW07] AXA-Winterthur: Jahresbericht 2007. AXA Winterthur, Winterthur 2007.
- [Ba00] Balzert, H.: Lehrbuch der Software-Technik – Software-Entwicklung. 2. edition, Spektrum Akademischer Verlag, Heidelberg, Berlin 2000.
- [Be01] Becker, J. et al.: Konstruktion von Methodiken – Vorschläge für eine begriffliche Grundlegung und domänenspezifische Anwendungsbeispiele. Institut für Wirtschaftsinformatik, Universität Münster, Münster 2001.
- [Br06] Brinkkemper, S.: Method engineering: engineering of information systems development methods and tools. In: Information and Software Technology 38 (1996) 4, S. 275-280.
- [De06] Dern, G.: Management von IT-Architekturen – Informationssysteme im Fokus von Architekturplanung und -entwicklung. Vieweg Verlag, Wiesbaden 2006.
- [De07] Deelmann, T.: Geschäftsmodellierung – Grundlagen, Konzeption und Integration. Logos, Berlin 2007.
- [Fa00] Farny, D.: Versicherungsbetriebslehre. VVW, Karlsruhe 2000.

- [FE07] FEAF: Federal Enterprise Architecture Framework, v.1.5. <http://www.whitehouse.gov/omb/egov/a-1-fea.html>.
- [FW07] Fischer, R.; Winter, R.: Ein hierarchischer Ansatz zur Unterstützung des IT/Business Alignment. In: Proceedings, eOrganisation: Service-, Prozess-, Market-Engineering, Karlsruhe 2007, S. 163-180.
- [Ge01] Gesamtverband der Deutschen Versicherungswirtschaft e.V. (GDV): Die Anwendungsarchitektur der deutschen Versicherungswirtschaft (VAA). <http://www.gdv-online.de/vaa/>.
- [Ha05] Hafner, M.: Entwicklung einer Methode für das Management der Informationssystemarchitektur im Unternehmen. Dissertation, Universität St. Gallen, Institut für Wirtschaftsinformatik, Difo-Druck, Bamberg 2005.
- [HR06] Huschens, J.; Rumpold-Preining, M.: IBM Insurance Application Architecture (IAA) – An overview of the Insurance Business Architecture. In: Bernus, P. M., K.; Schmidt, G. (Hrsg.): Handbook on Architectures of Information Systems. Springer, Berlin 2006, S. 669-692.
- [IE00] IEEE: IEEE Recommended Practice for Architectural Description of Software Intensive Systems (IEEE Std 1471-2000). New York, NY 2000.
- [Ke07] Keller, W.: IT-Unternehmensarchitektur – Von der Geschäftsstrategie zur optimalen IT-Unterstützung. Dpunkt Verlag 2007.
- [Ki07] Kilov, H.: Business Models – A Guide for Business and IT. Prentice Hall PTR, Upper Saddle River 2002.
- [KW07] Kurpuweit, S.; Winter, R.: Viewpoint-based Meta Model Engineering. In: Proceedings, 2nd International Workshop on Enterprise Modelling and Information Systems Architectures, St. Goar/Rhine, Germany 2007, S. 143-161.
- [KW92] Kumar, K.; Welke, R. J.: Methodology Engineering – A Proposal for Situation-specific Methodology Construction. In: Cotterman, W.; Senn, J. A. (Hrsg.): Challenges and Strategies for Research in Systems Development. John Wiley & Sons, New York 1992, S. 257-269.
- [La05] Lankhorst, M.: Enterprise Architecture at Work: Modelling, Communication and Analysis. Springer, Berlin et al. 2005.
- [Le04] Leist, S.: Methoden zur Unternehmensmodellierung – Vergleich, Anwendungen und Diskussionen der Integrationspotenziale. Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2004.
- [LV07] Leppänen, M.; Valtonen, K.; Pulkkinen, M.: Towards a Contingency Framework for Engineering an Enterprise Architecture Planning Method. In: Proceedings, 30th Information Systems Research Seminar in Scandinavia, Tampere, Finland 2007.
- [MZ07] Magalhaes, R.; Zacarias, M.; Tribolet, J.: Making Sense of Enterprise Architectures as Tools of Organizational Self-Awareness (OSA). In: Proceedings, Proceedings of the Second Workshop on Trends in Enterprise Architecture Research (TEAR 2007), June 6 2007, St. Gallen, Switzerland 2007, S. 61-70.
- [Ni05] Niemann, K. D.: Von der Unternehmensarchitektur zur IT-Governance: Leitfaden für effizientes und effektives IT-Management. (Edition CIO) edition, Vieweg 2005.

- [Oe07] Österle, H. et al.: Business Engineering – Core-Business-Metamodell. In: WISU – Das Wirtschaftsstudium 36 (2007) 2, S. 191-194.
- [Oe95] Österle, H.: Business Engineering: Prozess- und Systementwicklung. 2. Aufl. edition, Springer, Berlin et al. 1995.
- [Pa06] Pavlak, A.: Enterprise Architecture: Lessons from Classical Architecture. In: Journal of Enterprise Architecture 2 (2006) 2, S. 20-28.
- [Ro04] Rottwinkel, A.: Management von Partnerkontakten in Versicherungsunternehmen. Logos, Berlin 2004.
- [Sc05] Schekkerman, J.: The Economic Benefits of Enterprise Architecture: How to Quantify and Manage the Economic Value of Enterprise Architecture. Trafford Publishing 2005.
- [TL04] ter Doest, H.; Lankhorst, M.: Tool Support for Enterprise Architecture. Telematica Instituut, Enschede 2004.
- [To07] TheOpenGroup: The Open Group Architecture Framework TOGAF – 2007 Edition (Incorporating 8.1.1). Van Haren, Zaltbommel 2007.
- [VB06] Versteeg, G.; Bouwman, H.: Business architecture: A new paradigm to relate business strategy to ICT. In: Information Systems Frontiers 8 (2006) 2, S. 91-102.
- [We07] Weiss, D.: The Business Architecture Viewpoint: Defining the Bricks. Gartner Research 2007.
- [WF07] Winter, R.; Fischer, R.: Essential Layers, Artifacts, and Dependencies of Enterprise Architecture. In: Journal of Enterprise Architecture 3 (2007) 2, S. 7-18.
- [WM04] Whittle, R. M., Conrad B.: Enterprise Business Architecture. Auerbach Publications, Boca Raton 2004.
- [Za87] Zachman, J. A.: A Framework for Information Systems Architecture. In: IBM Systems Journal 26 (1987) 3, S. 276-292.

Ein Annotationsansatz zur Unterstützung einer ganzheitlichen Geschäftsanalyse

Sylvia Radeschütz, Florian Niedermann, Bernhard Mitschang

Institut für Parallele und Verteilte Systeme
Universität Stuttgart
Universitätsstr. 38
70569 Stuttgart

{sylvia.radeschuetz | bernhard.mitschang}@ipvs.uni-stuttgart.de
florian.niedermann@gmx.de

Abstract: Die Verbesserung der Geschäftsprozesse in einem Unternehmen spielt eine immer wichtigere Rolle, um Wettbewerbsvorteile gegenüber der Konkurrenz zu erlangen. Dafür ist eine umfassende Analyse nötig über alle verfügbaren Informationen in diesem Unternehmen. Aktuelle Verfahren konzentrieren sich entweder auf die Analyse von Prozessdaten oder die Analyse von operativen Anwendungsdaten, die typischerweise in einem Data Warehouse vorliegen. Für die Ausführung einer tiefergehenden Analyse ist es jedoch notwendig, Prozessdaten und operative Daten zu verknüpfen. Dieser Beitrag stellt zwei Ansätze vor, welche es ermöglichen, diese Daten effektiv und flexibel zusammenzuführen. Der erste Ansatz stellt eine direkte Verknüpfung von Entitäten aus den Prozessdaten mit Entitäten aus den operativen Daten her. Die Verknüpfung im zweiten Ansatz beruht hingegen auf der semantischen Beschreibung der Daten. Beide Methoden sind in einem Werkzeug realisiert.

1 Einleitung

Vor dem Hintergrund der Globalisierung und der Verbreitung des E-Business setzen Unternehmen vermehrt auf die Etablierung hochflexibler Geschäftsprozesse. Workflowsysteme automatisieren die Geschäftsprozesse und verbessern ihre Qualität. Dabei entscheidet eine schnelle Anpassung und Optimierung der Geschäftsprozesse an neue Marktsituationen oftmals über den Erfolg eines Unternehmens [We05]. Heutige Analyseverfahren im Unternehmen fokussieren sich entweder auf die Prozessanalyse (Monitoring und Workflow (WF) Mining) oder auf Analysen über operative Anwendungsdaten (Business Intelligence mit OLAP-Analysen und Data Mining). Zur weiteren Optimierung der Prozesse sind jedoch ganzheitliche Analysen notwendig, die Daten aus Prozessen und operativen Systemen gemeinsam auswerten. Diese sind bisher nur mit Hilfe aufwändiger manueller Aktivitäten zur Analysezeit möglich.

Somit verzichten viele Unternehmen bei der Geschäftssteuerung auf das Potential solcher umfassenden Analysen und führen die Analyse der Prozesse und der operativen Daten mit Hilfe getrennter Systemumgebungen durch. Die Prozessanalyse allein zeigt zwar auf, wie z.B. die Bearbeitungszeit der Aktivitäten durch bestimmte Angestellte variiert, aber sie gibt keinen Aufschluss darüber, welche Faktoren (z.B. Arbeitserfahrung oder Fortbildung) dabei eine Rolle spielen. Die Daten für diese Faktoren liegen in einem operativen Data Warehouse (DWH). Auch eine rein operative Datenanalyse führt nicht zu einer Prozessoptimierung, da im DWH wiederum wichtige Prozessdaten für die Analyse fehlen.

Werden Prozessdaten und operative Daten jedoch während des ETL (Extraktion, Transformation, Laden) zusammengeführt (Matching in Abb. 1) und in einem DWH integriert, können ganzheitliche, unternehmensweite Analysen einen wertvollen Beitrag zu weiteren Prozessoptimierungen leisten, indem Faktoren aus beiden Datenquellen berücksichtigt werden. Prozesseigenschaften wie die Prozesslaufzeit können nun bezüglich ihrer Kausalitäten organisiert werden. So kann z.B. geprüft werden, welchen Einfluss Arbeitserfahrung und Fortbildungen auf die Bearbeitungszeit bestimmter Prozessaktivitäten haben.

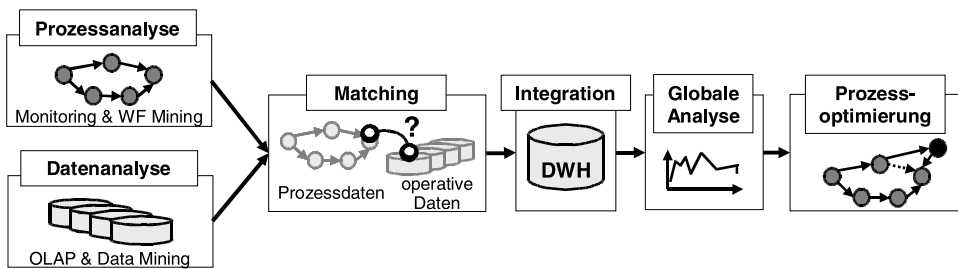


Abbildung 1: Unternehmensweite Geschäftsanalyse

In diesem Beitrag werden zwei Match-Methoden vorgestellt, die es ermöglichen, diese Daten sowohl direkt als auch auf semantischer Ebene zu verknüpfen. Diese Methoden sind Teil eines größeren Frameworks, das zunächst in Kapitel 2 eingeführt wird. Basierend auf den Methoden wird ein Editor entwickelt. Die Match-Methoden selbst, wie auch die Konzepte und die Realisierung des Editors, werden in Kapitel 3 und 4 beschrieben. In Kapitel 5 folgen weiterführende Betrachtungen, wie eine ganzheitliche Analyse mit Hilfe der daraus resultierenden Verknüpfung der Daten durchgeführt werden kann. Die Methoden und der Editor werden in Kapitel 6 anschließend evaluiert. Verwandte Arbeiten (Kapitel 7) und eine Zusammenfassung (Kapitel 8) bilden den Abschluss.

2 Matching Framework

Wie in Kapitel 1 dargelegt, ist es für eine umfassende Optimierung der Geschäftsprozesse notwendig, dass alle relevanten Prozessdaten und operativen Daten zusammengeführt werden. Die Verknüpfung der Daten kann manuell oder (halb)automatisch erfolgen. Abb. 2 gibt eine Übersicht über unser Framework, das Match-Methoden über verschie-

dene Workflow-Management-Systeme, operative Datenquellen und semantische Beschreibungssprachen hinweg integriert und anschließend den Match in einer Annotati-
onstabelle ablegt.

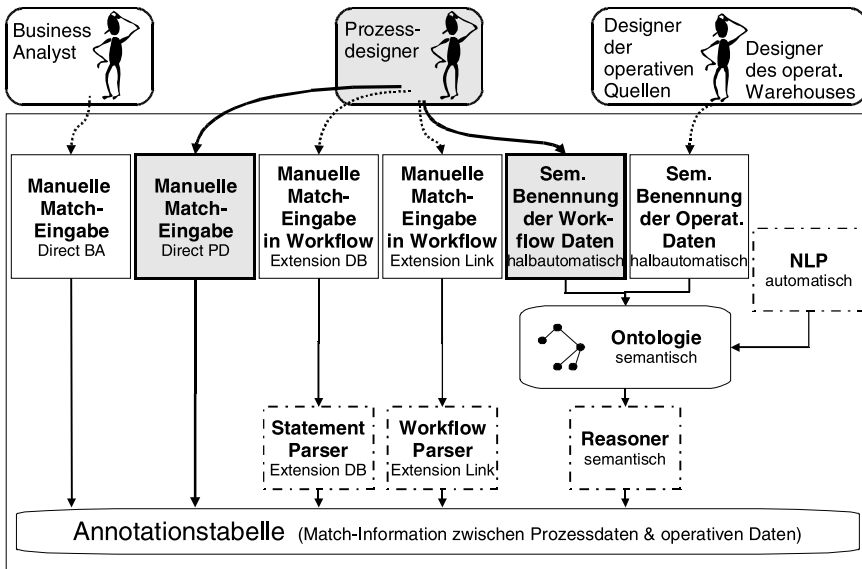


Abbildung 2: Architektur des Matching Frameworks

Das Framework bietet mehrere Alternativen für eine passende Match-Methode je nach Situation, z.B. abhängig davon, ob die genaue Verknüpfung dem Prozessdesigner bekannt ist oder er nur die semantische Bedeutung der Daten kennt. Die oberen Rechtecke im Framework in Abb. 2 zeigen den Namen der Methode und die verschiedenen Eingaben der Personen, die am Match beteiligt sind und auf die im Folgenden näher eingegangen wird. Grau hinterlegt sind die beiden Match-Methoden des Prozessdesigners (Direct PD und halbautomatisch), die durch den Annotationseditor, der in diesem Beitrag vorgestellt wird, realisiert und unterstützt werden. Eine allgemeine Übersicht über das gesamte Framework mit allen Methoden findet man in [RML08]. Gegebenenfalls sind für die Ausführung der Methoden im Framework neben den Benutzereingaben zusätzliche Tools nötig. Diese sind in Abb. 2 in gestrichelten Boxen dargestellt.

2.1 Business Analyst

Die aufwändigste Möglichkeit für eine direkte Verknüpfung ist ein manueller Match durch den Business Analysten (Direct BA). Hier müssen passende Matches zur Analysezeit gefunden werden, was aufgrund der riesigen Datenmengen im operativen DWH und im Audit Trail, in dem die Prozessdaten protokolliert werden, sehr zeitintensiv und fehleranfällig ist. Des Weiteren muss ein Matching der Daten gegebenenfalls bei jeder Analyse neu durchgeführt werden. Daher sind andere Methoden (siehe Abb. 2) zu bevorzugen.

2.2 Prozessdesigner

Dem Prozessdesigner stehen im Framework vier Alternativen zur Verfügung, die ihn beim Matching der Daten unterstützen. Er kann die konkrete Verknüpfung direkt angeben und abspeichern, wenn der Prozess fertig gestellt ist (Direct PD). Dazu gibt er zu jeder Prozessvariablen die passende operative Datenquelle an. In den nächsten Kapiteln wird diese Methode als Teil des Editors im Detail beschrieben.

Es kann aber auch sein, dass der Match schon Teil des Workflows ist. Dies ist der Fall, wenn der Workflow selbst um Anfragen auf operative Datenquellen erweitert wird, auf die sich der Workflow bezieht (Extension DB), weil die Anfrageergebnisse im weiteren Workflow benötigt und weiter verarbeitet werden müssen. Der Prozessdesigner muss also schon während des Prozessdesigns Datenbank-Statements entweder direkt im Workflowcode (z.B. IBM BPEL/SQL PlugIn [IBM07]) oder in einen Datenbankadapter (z.B. Oracle Database Adapter [VD07]) im Workflow einbetten, wodurch eine Verbindung zur jeweiligen operativen Datenquelle hergestellt wird. Der Match wird später durch einen Parser extrahiert.

Während des Prozessdesigns wird auch die nächste Match-Methode angewandt. Hier ergänzt der Prozessdesigner jede Prozessvariable im Prozess durch ein Attribut, das auf die betreffende operative Datenquelle verweist (Extension Link). Diese Methode verlangt jedoch eine Erweiterung der Workflowengine, die diese Attributerweiterung bei dem Deployment (Kompilieren des Workflows) erkennt und im Audit Trail abspeichert.

Die nächste Match-Alternative (halbautomatisch) eignet sich für die Annotation von Prozessvariablen, zu denen dem Prozessdesigner die entsprechende operative Datenquelle nicht bekannt ist. Er gibt zu jeder Variable die Referenz zu der passenden semantischen Beschreibung in einer Ontologie an. Damit die dazugehörige Datenquelle später automatisch gefunden werden kann, muss jedoch auch diese entsprechend semantisch annotiert sein. Das ist die Aufgabe des DWH-Designers und wird in Kapitel 2.3 behandelt.

2.3 Designer des Operativen Data Warehouses

Für eine halbautomatische Zusammenführung annotiert der Designer der operativen Datenquellen oder der DWH-Designer die Daten mit Hilfe einer Ontologie semantisch. Ein Reasoning Tool kann anschließend auf Basis der Annotationen und der Ontologie die Prozessdaten mit den operativen Daten matchen.

Für eine komplett automatische Verknüpfung wird ein Natural Language Processing Tool (NLP) benötigt, das in einer eingeschränkten Domäne zu jeder Variable und operativen Datenquelle abhängig von deren Namen das passende semantische Konzept findet. Die Annotation läuft so voll automatisch ab ohne Prozessdesigner und DWH-Designer.

3 Annotationsmechanismus

Um den manuellen Ansatz „Direct PD“ und den halbautomatischen semantischen Ansatz unseres Frameworks umzusetzen, ist ein Mechanismus zur Annotation der Workflowmodelle notwendig.

Für die Auswahl eines Mechanismus legen wir zu Grunde, dass die Workflowmodelle in der Sprache BPEL [Jo07] beschrieben und auf sogenannten BPEL-Engines ausgeführt werden. Diese Annahme ist plausibel, da BPEL sich über die letzten Jahre sowohl auf Anwender- als auch auf Herstellerseite als dominierender Standard für die Beschreibung von Workflows etabliert hat.

3.1 Annotierbare Elemente

Für die Auswahl eines Annotationsmechanismus muss zunächst geklärt werden, welche Daten überhaupt annotiert werden sollen. Ausgangsbasis sind dabei die Daten, die während der Ausführung des Workflows in die sogenannte Auditdatenbank geschrieben werden. Die Standardisierung der Datenformate in diesem Bereich (siehe [ZM04, Wm98]) ist noch nicht sehr weit fortgeschritten bzw. umgesetzt. Dennoch enthalten sämtliche Auditdatenbanken der von uns untersuchten BPEL-Engines für jedes Ereignis im Prozess (z.B. Aufruf eines Web Services, Aktualisierung einer Variable) zumindest die folgenden Daten:

- (a) ID des Workflowmodells
- (b) ID der Workflowinstanz
- (c) Typ und die ID des Ereignisses
- (d) Ein Zeitstempel für den Zeitpunkt des Ereignisses
- (e) Name und Typ der BPEL-Aktivität, welche das Ereignis ausgelöst hat: Wenn die Aktivität einen Web Service aufgerufen hat, wird zusätzlich eine ID des Services angegeben.
- (f) Name sowie der (neue) Wert der Variablen, wenn das Ereignis eine Variable verändert hat.

Die folgenden zwei Kriterien sind für die Auswahl der zu annotierenden Elemente zu beachten: Das erste Kriterium ist die *Verknüpfbarkeit*. Ziel unseres Frameworks ist die Verknüpfung von operativen Datenquellen und Prozessdaten. Eine Annotation ist daher nur für die Elemente sinnvoll, welche in irgendeiner Form sowohl in den operativen Daten als auch auf Prozessseite vorkommen. Das zweite Kriterium ist die *Nützlichkeit*. Ein Element ist im Sinne unserer Analyse nützlich, wenn es zusätzliche Einsichten und Erkenntnisse aus Geschäftssicht ermöglicht. Rein technische Steuerungsinformationen, wie sie z.B. für das Routing von Nachrichten zwischen Services verwendet werden, sind daher weniger interessant.

Die Anwendung beider Kriterien ergibt, dass sämtliche Elemente bis auf die Variablen von der Annotation ausgeschlossen werden können. Im Wesentlichen geht dies auf das Verknüpfbarkeitskriterium zurück, da alle anderen Elemente auf Seite der operativen Datenquellen kein entsprechendes Gegenstück aufweisen.

Variablen sind die Datencontainer eines BPEL-Prozesses. Sie bilden Input- und Outputdaten des Prozesses sowie mögliche Zwischenergebnisse ab. Ihre Struktur entspricht entweder einem XML Schema Type oder einer WSDL Message. XML Schema ist ein Standard zur Beschreibung der Datenstruktur von XML-Dokumenten. WSDL Messages sind in WSDL 1.1 definierte Nachrichten, die von und mit Web Services ausgetauscht werden. Die Annotation der Variablen kann daher auf drei verschiedenen Ebenen erfolgen

(siehe Abb 3):

Man kann die Variable „InputVariable“ in (1) selbst annotieren oder Teile einer Message („parameters“ in (2)). Falls die Variable komplex mit XML Schema Types (3) aufgebaut ist, ist ebenso eine Annotierung der einzelnen Elemente, z.B. der „EmployeeID“ möglich. Wir haben uns dafür entschieden, es dem Benutzer zu ermöglichen, die Variablen bis auf die Ebene der XML-Schemastruktur zu annotieren. Die Annotation der Variable selbst ist zwar einfach umzusetzen, bringt aber wenig neue Erkenntnisse. Ähnliches gilt für die Message Parts. Hier kommt das Problem hinzu, dass das Message-Konzept mit der Einführung von WSDL 2.0 entfernt wurde [Ch07].

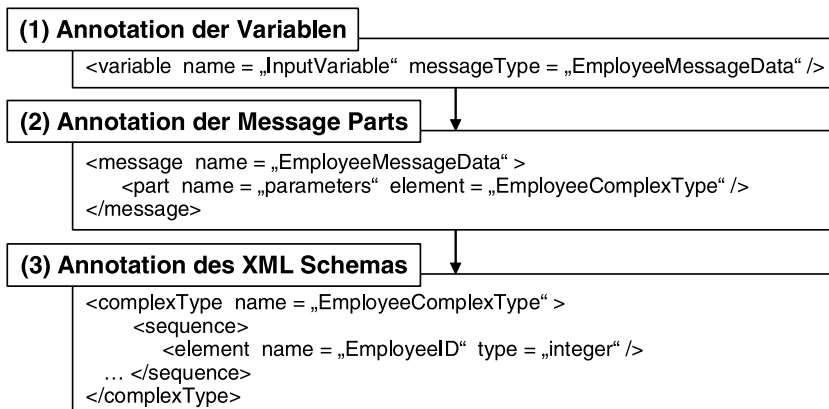


Abbildung 3: Annotationstiefe bei Variablen

3.2 SAWSDL

Für die Annotation von Web Services und ihrer Daten haben sich in den letzten Jahren im Rahmen des gestiegenen Interesses am Semantic Web eine Reihe von Standards etabliert. Auf der einen Seite stehen „heavy-weight“ Ansätze, die eine komplett neue konzeptuelle semantische Repräsentation eines Web Service bereitstellen. Die prominentesten Vertreter dafür sind WSM (Web Service Modeling Language) [Br05] und OWL-S (Web Ontology Language for Web Services) [Ma04].

Auf der anderen Seite stehen Standards wie SAWSDL (Semantic Annotations for WSDL) [FL07], die lediglich die Beschreibungssprache WSDL um zusätzliche Elemente erweitern. Dieser Ansatz ist für unser Framework besser geeignet als z.B. WSML oder OWL-S. Wir wollen fokussiert bestimmte Elemente einer WSDL- oder XML-Schema-Datei beschreiben und nicht die Semantik eines gesamten Services erfassen.

Daher werden wir im Folgenden auf die verschiedenen Elemente von SAWSDL und ihrer Verwendung für die Umsetzung der direkten und semantischen Datenverknüpfung eingehen. Attribute, die nicht Teil des SAWSDL-Standards sind und von uns eingeführt werden, sind mit einem „*“ markiert.

Semantische Verknüpfung

- **modelReference**: Das „modelReference“-Attribut wird verwendet, um XML-Elemente mit Konzepten einer Ontologie zu verknüpfen. SAWSDL schreibt keine spezielle Struktur für diese Referenz vor, es hat sich aber die folgende Konvention eingebürgert, um das Konzept in der Ontologie zu adressieren:

<Ontologie ID> „#“ <Konzept ID>

- **liftingSchemaMapping** und **loweringSchemaMapping**: Diese Attribute ermöglichen es, eine Datei zu referenzieren, die die Transformation von XML in die Ontologiesprache (liftingSchemaMapping) bzw. in die andere Richtung (loweringSchemaMapping) realisiert.
- **modelType***: Zur Unterscheidung von verschiedenen Ontologiesprachen haben wir dieses Attribut eingeführt. Es vereinfacht die Arbeit eines Parsers bei dem automatischen Matching und verbessert die Lesbarkeit.

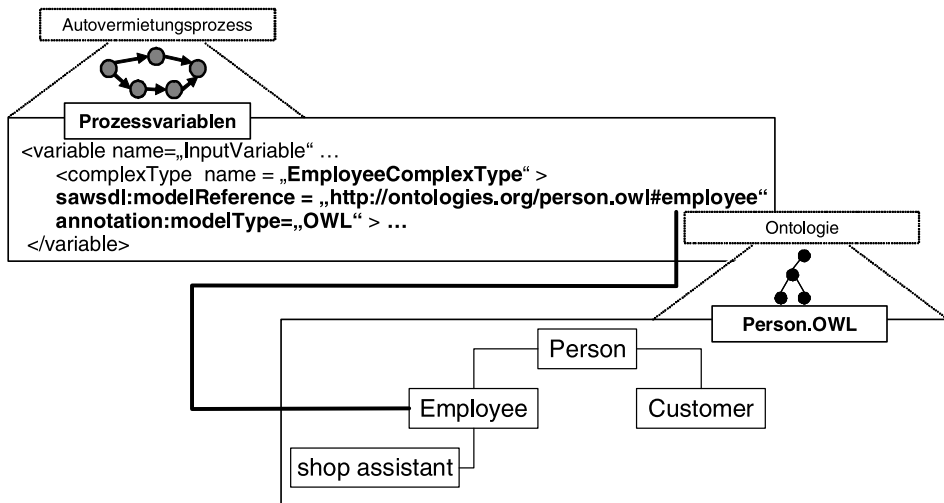


Abbildung 4: SAWSDL für semantische Verknüpfung

4 Annotationseditor

Zur Umsetzung der Annotationsmethode stellen wir einen Editor bereit. Der Editor unterstützt einen Prozessdesigner dabei, BPEL-Prozesse zu laden, zu annotieren und die annotierten Daten in angemessener Form abulegen. Des Weiteren ermöglicht er es, XML-Schema- und WSDL-Dokumente direkt in ihrer Datei zu annotieren und abzuspeichern. Wie Abb. 6 zeigt, orientiert sich der Aufbau des Editors an gängigen Entwicklungsumgebungen und ähnlichen Werkzeugen.

Der Projektexplorer (1) zeigt die aktuell geladenen Prozesse und Dokumente an. Wie in der Abbildung zu sehen ist, ist aktuell nur der BPEL-Autovermietungsprozess „Rental-Process“ in den Workspace des Editors geladen.

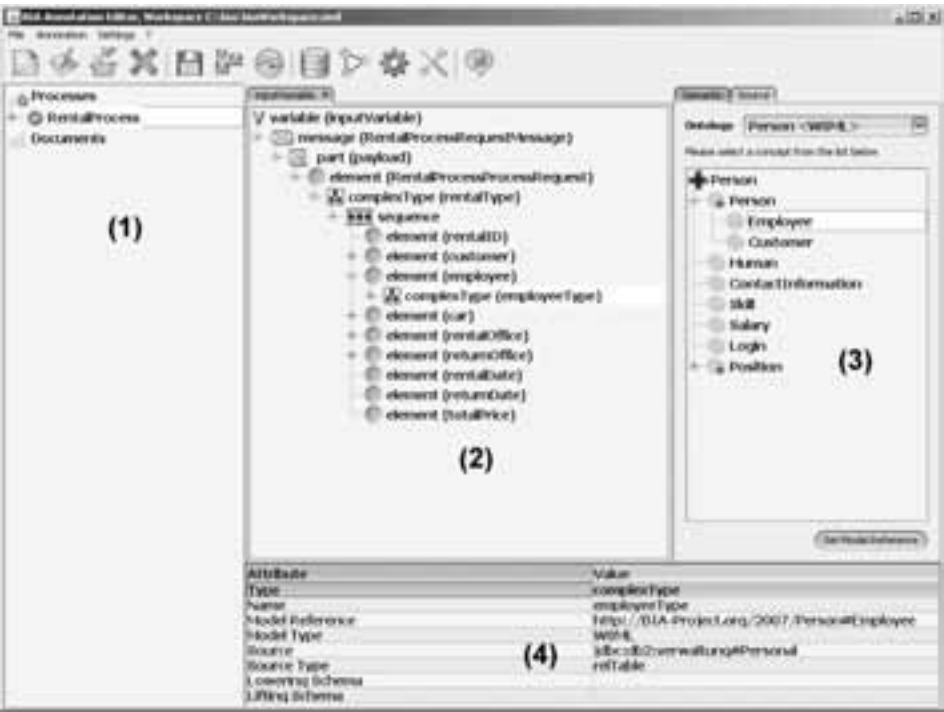


Abbildung 6: Screenshot des Editors

Das zentrale Element des Editors ist der Annotationsexplorer (2), welcher die Struktur von BPEL-Variablen sowie WSDL- und XML-Schema-Dateien visualisiert. In der Abbildung ist die BPEL-Variable „InputVariable“ geöffnet und das Element „employee“ ausgewählt. Der Benutzer kann einzelne Variablenelemente auswählen und entweder über die Attributstabelle (4) oder den graphischen Editor (3) für Ontologie- und Datenverknüpfung annotieren. Die Attributstabelle zeigt die textuellen Werte der Annotationsattribute an. Es ist möglich, mit Hilfe der Tabelle die Werte der Annotationsattribute manuell zu verändern. Meist ist es jedoch einfacher, dafür den graphischen Editor zu verwenden.

Der Editor ist unter Java Version 1.5 implementiert. Zur Erbringung seiner Funktionalität verwendet er eine Reihe von APIs wie z.B. WSDL4J [Ws06] bzw. Apache Woden [Ap07] für das Laden der WSDL1.1- bzw. der WSDL2.0-Beschreibungen oder Jena [Je07] bzw. WSMO4J [Ws07] für das Laden und Visualisieren der OWL- bzw. WSML-Ontologien.

Unterstützt werden BPEL-Prozesse für IBM Process Server [IBM07] und Oracle SOA Suite [Or07]. Die Architektur des Editors bietet Erweiterungspunkte, die die Unterstützung weiterer Engines und weiterer operativer Datenquellen, wie z.B. XML, relativ problemlos zu einem späteren Zeitpunkt erlauben.

5 Ganzheitliche Analyse in der Praxis

Dieses Kapitel zeigt anhand eines Fallbeispiels, wie der Editor für das Zusammenführen von Daten verwendet wird, um später eine ganzheitliche Analyse im Unternehmen durchführen zu können. In dem Praxisbeispiel will eine Autovermietungsfirma ihre Geschäftsprozesse optimieren. Dazu werden hier die Prozesse zur Vermietung eines Autos an einen Kunden herausgegriffen und betrachtet. Viele Prozessdaten sind für die Analyse von Belang, uns interessieren aber besonders die Prozessvariablen, die mit Daten aus operativen Systemen verknüpft werden, um durch die Analyse zusätzliche Informationen zu generieren. Der Annotationseditor erleichtert die Zusammenführung, die in Abb. 7 schematisch dargestellt ist.

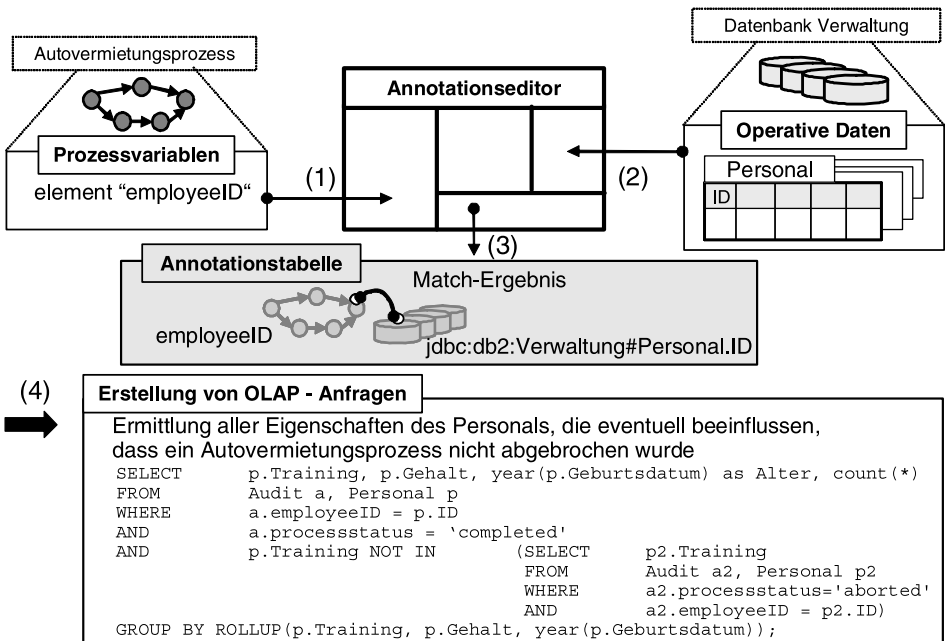


Abbildung 7: Ganzheitliche Analyse basierend auf direktem Match

In (1) werden die Variablen aus dem BPEL-Autovermietungsprozess, u.a. auch die „EmployeeID“ des Angestellten, der die Autovermietung abgeschlossen hat, in den Annotationseditor geladen. Operative Daten aus anderen Systemen, wie hier aus der relationalen Verwaltungsdatenbank mit der Tabelle „Personal“ aus der Personalabteilung werden in (2) in den Editor übertragen. Der Prozessdesigner wählt nun das Element „EmployeeID“ der Prozessvariablen „InputVariable“ aus und verknüpft es mit der Spalte ID aus der Tabelle Personal. Die Verknüpfung wird in (3) als Match-Ergebnis in eine Annotationstabelle abgelegt.

Diese Match-Tabelle erleichtert es dem Analysten später, die zusammengehörigen Verbindungen zu finden, sodass er die OLAP-Anfrage zügiger erstellen kann. Wenn der Analyst die OLAP-Anfrage erstellt (4), kann er im Match-Ergebnis nachschauen, welche Variablen zu welchen operativen Daten passen und für eine ganzheitliche Analyse miteinander verbunden werden müssen (Join-Operator). Mit der OLAP-Anfrage können z.B. die Eigenschaften der Angestellten ermittelt werden, die dazu führen, dass ein Mietprozess erfolgreich abgewickelt wird und die im Gegensatz dazu jenem Personal fehlen, deren bearbeitete Prozesse abgebrochen werden. In diesem Beispiel werden hierfür die Eigenschaften Training, Gehalt und Alter der Angestellten untersucht und mit den Prozessdaten aus dem Audit Trail verknüpft, um den Prozessstatus zu erhalten. Der Audit ist hier beispielhaft als eine Tabelle dargestellt. Abhängig vom Workflow-Management-System und seiner Protokollierung ist dies evtl. eine komplexe Subquery. Zur Analyse werden die Personaleigenschaften in verschiedenen Gruppierungen dargestellt (Rollup-Operator). Unsere zukünftigen Forschungsarbeiten beschäftigen sich ausführlicher mit diesen OLAP-Anfragen.

Gegebenenfalls wird nun aus der Analyse ersichtlich, dass die Eigenschaft, psychologisch geschult zu sein, ein positives Ergebnis auf den Ablauf der Mietprozesse hat. Eine Optimierung der Geschäftsprozesse erfolgt dann dadurch, dass die Aktivitäten zur Kundenberatung im Mietprozess hauptsächlich von psychologisch trainierten Mitarbeitern ausgeführt werden und zudem eine Psychologie-Fortbildung auch den anderen Kollegen angeboten wird.

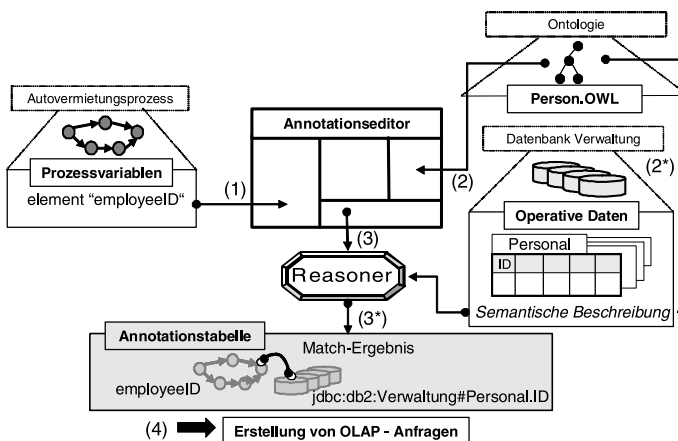


Abbildung 8: Ganzheitliche Analyse basierend auf semantischem Match

Für die semantische Annotation in Abb. 8 wird in (2) statt der operativen Daten, eine Ontologie geladen. So wird durch den Editor die Variable mit einem semantischen Konzept versehen. Das komplexe Element „Employee“ der Variable „InputVariable“ referenziert das OWL-Konzept „Employee“ und wird so in (3) in die Annotationstabelle abgelegt. Auch die operative Datenquelle muss mit einem Editor semantisch annotiert werden (2*). Schließlich findet ein Reasoner automatisch die passenden Verknüpfungen, die in (3*) als Match-Ergebnis abgespeichert werden. So kann auch mit Hilfe dieses Matches in (4) eine OLAP-Anfrage zwecks Analyse formuliert werden.

6 Evaluation

Dieses Kapitel analysiert den Nutzen einer Verknüpfung zwischen Prozessvariablen und operativen Daten und inwiefern der Annotationseditor dabei nützlich ist. Dazu werden zunächst die hier verwendeten Match-Methoden im Allgemeinen und danach der Editor selbst evaluiert. Da wir zu einem so frühen Zeitpunkt noch über keine empirischen Studien verfügen, stützt sich die Evaluierung auf eine von uns entwickelte Fallstudie und qualitative Überlegungen. Anschließend erfolgt eine Analyse des betriebswirtschaftlichen Nutzens des Editors und der Match-Methoden.

6.1 Evaluation der Match-Methoden

In den vorherigen Kapiteln wurden beide Match-Methoden, das direkte Matching und das semantische Matching, ausführlich erläutert. Sie werden in unterschiedlichen Situationen verwendet. Je nach Fachwissen des Designers ist die direkte oder semantische Methode schneller und effektiver. Das direkte Verknüpfen der Prozessdaten mit operativen Daten funktioniert sehr effizient, wenn der Prozessdesigner sich auch bei den operativen Datenquellen auskennt. Dies ist nützlich bei einem Match zwischen Entitäten, deren Daten ähnlich strukturiert sind, z.B. um ihre Schlüsselwerte miteinander zu verbinden (siehe Abb. 8 aus Kapitel 5).

Sobald komplexere Transformationen für das Matching der Daten notwendig sind oder dem Designer die zu matchenden operativen Datenquellen unbekannt sind, ist das semantische Matching effektiv. Dafür müssen jedoch auch die operativen Datenquellen von dem operativen DWH-Designer semantisch annotiert vorliegen. Außerdem ist das Vorhandensein eines Reasoners und einer Ontologie erforderlich, in deren Domäne sich der Designer für eine genaue Annotation der Daten gut auskennen muss. Das Matching der Daten ist auf diese Domäne beschränkt im Gegensatz zum direkten Matching, bei dem die Daten vom Designer aus verschiedenen Domänen verknüpft werden können. Durch die Bereitstellung dieser Ontologie entsteht zwar ein zusätzlicher Aufwand. Dieser fällt vor dem Hintergrund aktueller Entwicklungen jedoch nicht sehr stark ins Gewicht. Der Trend geht nämlich in die Richtung semantischer Web Services [CS06]. Daher bekommen entsprechende semantische Technologien gerade großen Auftrieb. Semantische Annotationen mit SAWSDL werden standardisiert und für die Analyse wieder verwertbar.

6.2 Evaluation des Editors

Die Annotationen für die beiden Match-Methoden können mit dem Editor effektiv umgesetzt werden. Der Editor befolgt die verschiedenen Richtlinien für eine graphische Schnittstelle zur Mensch-Computer-Interaktion aus dem ISO Standard [ISO06] und erleichtert so dem Benutzer seine Bedienung:

- *Selbstbeschreibungsfähigkeit*: Der Editor ist selbsterklärend. Der Benutzer erhält zudem Informationen zu Icons mit Hilfe von Tooltips und zusätzliche Textinformation während der Ausführung längerer Dialogfunktionen. Operative Datenquellen bzw. Ontologien werden intuitiv als Baumstruktur dargestellt. Alle Funktionen sind konsistent aufgebaut.
- *Aufgabenangemessenheit*: Nur für die Annotation wichtige Informationen werden im Editor dargestellt. Dadurch wird eine Reizüberflutung des Benutzers vermieden.
- *Steuerbarkeit*: Der Benutzer hat die Kontrolle über den Editor und kann jederzeit eine Funktion abbrechen.
- *Individualisierbarkeit*: Bestimmte Einstellungen sind durch den Benutzer konfigurierbar. Dazu gehören z.B. das automatische Laden von Ontologien und Abspeichern der Annotationen. Außerdem können die Teile der Prozessvariablen angegeben werden, die man im Editor annotieren darf.
- *Fehlerrobustheit*: Fehlerhafte Eingaben werden möglichst vermeidbar gemacht über die Dialoggeschlossenheit der Funktionen und das Deaktivieren von Funktionen, die zu dem jeweiligen Zeitpunkt nicht ausführbar sind.
- *Erlernbarkeit*: Zwar ist die Benutzung des Editors insgesamt selbsterklärend, als zusätzliche Unterstützung sind aber einige Tutorials vorhanden.
- *Erwartungskonformität*: Der Editor ist auf die Bedürfnisse des Prozessdesigners für das Matching der Daten ausgerichtet.

6.3 Betriebswirtschaftliche Nutzenpotentiale

Abb. 9 fasst die mit der Annotationsmethode und dem Ausführen einer ganzheitlichen Analyse verfolgten Nutzenpotentiale zusammen. Das ganzheitliche Analyseverfahren ermöglicht es einem Unternehmen, die Effizienz und Effektivität seiner Prozesse zu steigern. Indem die Annotationsmethode und der Editor Kosten-, Zeit- und Qualitätseffekte positiv beeinflussen, wird der Einsatz des Analyseverfahrens erst praktikabel („Enabler“).

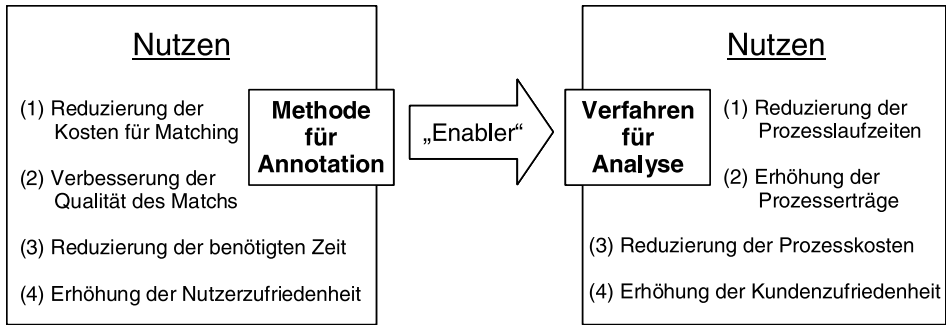


Abbildung 9: Nutzen der Annotationsmethode und der ganzheitlichen Analyse

Im Folgenden werden die Nutzenpotentiale der Methode und des Editors im Detail und analog zu Abb. 9 vorgestellt, die sich aus seiner Verwendung durch den Prozessdesigner gegenüber einer manuellen Annotierung zur Analysezeit oder ohne Editor ergeben:

(1) *Reduzierung der Kosten für Matching*: Der Aufwand und damit die Kosten des Personals für die Erstellung der Verknüpfung verringern sich durch die gut bedienbare graphische Benutzeroberfläche. Zudem muss der Match für die Variablen nur einmal pro Workflowmodell erfolgen und nicht für jede Analyse aufs Neue. Durch die Verwendung von SAWSDL im Editor ergeben sich noch weitere Kostenreduzierungen, da bereits erstellte Annotationen im Bereich der semantischen Web Services für die Analyse wieder verwendet werden können. Außerdem müssen bereits früher mit dem Editor erstellte semantische Annotationen nicht an neue operative Datenquellen angepasst werden, woraus sich weitere Einsparungen ergeben. Die Kosten zur Analysezeit verringern sich durch die Verwendung der direkten und semantischen Matches erheblich.

(2) *Verbesserung der Qualität des Matches*: Die Fehlerquote der direkten Verknüpfung zwischen Prozessdaten und operativen Daten sinkt. Bei dem semantischen Match können zudem Beziehungen gefunden werden, die einem Menschen nicht sofort aufgefallen wären. Eine standardisierte Ausgabe des Matches durch den Editor ermöglicht eine effektive Weiterverarbeitung durch den Analysten. Weiterhin erhöht die plattformübergreifende Nutzung verschiedener Workflow-Management-Systeme, Ontologiesprachen und beliebiger operativer Datenbanken die Qualität des Matches.

(3) *Reduzierung der benötigten Zeit*: Alle Faktoren, die bei den vorherigen Punkten eine Rolle spielen, können auch für die Zeitersparnis angeführt werden. Die graphische Benutzeroberfläche verkürzt die Zeit, die für eine Annotation benötigt wird. Die Analysten wiederum können effizienter und effektiver ihre Analysen erstellen, da sie bereits eine qualitativ wertvolle und standardisierte Verknüpfung vorliegen haben und nicht selbst danach suchen müssen. Dadurch ergibt sich insgesamt eine schnellere Analyse der Prozesse.

(4) *Erhöhung der Nutzerzufriedenheit*: Eine gute Handhabung des Editors (siehe auch Kapitel 6.2) und ein schnelles und zuverlässiges Match-Ergebnis erhöhen die Bereitschaft des Nutzers diesen Editor zu verwenden.

Die Nutzenfaktoren Zeit- und Kostenersparnis machen es sinnvoll, den Editor für die Verknüpfung zu verwenden. Die Erhöhung der Nutzerzufriedenheit führt dazu, dass der Editor auch wirklich benutzt wird und letztendlich die Annotationen für einen Match überhaupt erst erstellt werden, was ohne den Editor selten der Fall wäre. Ansonsten bliebe die Erstellung des Matches die Aufgabe der Analysten, was aufgrund des hohen Zeitaufwandes in der Regel nicht durchgeführt wird. Der Editor macht also ein ganzheitliches Analyseverfahren erst möglich.

Das Nutzenpotential, das sich aus dem Analyseverfahren und der daraus resultierenden Prozessoptimierung ergibt, ist nachfolgend aufgelistet:

(1) *Reduzierung der Prozesslaufzeiten:* Durch die Verknüpfung der Prozessdaten mit operativen Daten können Faktoren erkannt werden, die die Laufzeit eines Prozesses negativ beeinflussen. Mit der Erstellung neuer Prozessregeln, die diese Eigenschaften umgehen oder alternativ bearbeiten, kann die Laufzeit gesteigert werden.

(2) *Erhöhung der Prozesserträge:* Prozesse, die abgebrochen wurden oder nur wenig Gewinn einbringen, werden in Verbindung mit den operativen Datenquellen ganzheitlich analysiert und optimiert.

(3) *Reduzierung der Prozesskosten:* Durch die Reduzierung der Prozesslaufzeit werden oft auch die Prozesskosten reduziert. So können z.B. die Kosten der im Prozess eingesetzten Ressourcen wie Maschinen und Mitarbeiter optimiert werden.

(4) *Erhöhung der Kundenzufriedenheit:* Die ganzheitliche Analyse erlaubt es, Kundenbedürfnisse und -wünsche schneller zu erkennen und in die Geschäftsprozesse zu integrieren. Dies erhöht die Kundenzufriedenheit und kann Kunden langfristig an das Unternehmen binden.

7 Verwandte Arbeiten

In der Welt der Datenbanken liegen Methoden zur Informationsintegration im Fokus der Forschung. Dies betrifft sowohl die direkte Zusammenführung von Schemata [LN06] als auch das Matching auf Basis der semantischen Beschreibung der Datenquellen [DH05] mit dem Ziel, einen einheitlichen Zugriff auf unterschiedliche operative Datenquellen zu schaffen. Auch unsere Methoden zielen auf eine Schemaintegration ab. Jedoch betrachten wir eine andere Ebene der Integration. Die operativen Daten sollen mit Variablen gematcht werden, die von Workflows verwendet werden, also in einem ganz anderen Umfeld vorkommen als die operativen Daten. Zudem unterscheiden sich der Inhalt und die Struktur eines Audit Trails, in dem die Ablaufdaten des Workflows gespeichert werden, grundlegend von den operativen Datenquellen, was bei der Zusammenführung ebenso eine Rolle spielt.

Auf der Seite der Geschäftsprozesse beschreiben [CS06] und [He05] semantische Technologien mit dem Ziel einer automatischen Web-Service-Komposition. Dazu muss der Web Service als Ganzes beschrieben werden, während in unserer Arbeit vor allem eine detaillierte semantische Variablenannotation von Belang ist vor dem Hintergrund einer ganzheitlichen Analyse.

Mit WSMO Studio [Ws08] oder Radiant aus dem Meteor-S Projekt der Universität Georgia [Me05] existieren Editoren, die die semantische Annotation von WSDL-Dokumenten unterstützen. Beide Werkzeuge sind an bestimmte Ontologien gekoppelt, entweder an WSML-Ontologien im ersten Tool oder an OWL im zweiten. Die jeweils andere Ontologie kann graphisch nicht eingebunden werden. Auch IBM redet von einem semantischen PlugIn [IBM06] für den WebSphere Integration Developer. Die Benutzeroberfläche für die Annotation ist dann jedoch nur textuell und nicht graphisch ausführbar. Bei allen Editoren fehlt außerdem die direkte Integration mit operativen Datenquellen.

Für die direkte Verknüpfung von Prozessdaten und operativen Daten zur Prozessdesignzeit steht bisher außer unserem Editor mit der Methode „Direct PD“ keine Annotationsmethode zur Verfügung. Vorhandene Ansätze führen die Daten erst zur Analysezeit manuell zusammen [ZM04; Ca07]. Ein weiterer Ansatz von [ZM04] fügt künstliche Aktivitäten in den Workflow ein, die dann das Matching übernehmen. Die zusätzlichen Aktivitäten verfälschen jedoch die Semantik und evtl. sogar die Ausführungszeit der Prozesse, da sie das Workflowmodell verändern.

8 Zusammenfassung und Ausblick

Dieser Beitrag stellt zwei Methoden vor, wie Prozessdaten und operative Daten zusammengeführt werden können, um eine ganzheitliche Unternehmensanalyse zu ermöglichen. Die zwei Ansätze erlauben es, BPEL-Variablen mit einer Erweiterung von SAWSDL zu annotieren. Die Annotation zeigt auf die operative Datenquelle, die sich auf dasselbe reale Objekt bezieht wie die Prozessvariable. Der Link kann vom Prozessdesigner entweder direkt spezifiziert werden oder er referenziert ein Konzept aus einer Ontologie, das diese Variable semantisch beschreibt, um den Match später durch einen Reasoner herzustellen.

Die Evaluierung der beiden Methoden und ihre Betrachtung in einem globalen Fallbeispiel haben gezeigt, dass beide Methoden je nach Situation, in der sie eingesetzt werden, sehr hilfreich für den Prozessdesigner sind und einen großen Nutzen für die Optimierung der Geschäftsprozesse und damit für ein Unternehmen erzielen. Mit Hilfe eines Editors hat der Prozessdesigner die Möglichkeit, die Methoden effektiv und unter minimalem Lernaufwand anzuwenden. Bisher können damit BPEL-Variablen für die BPEL-Engines IBM Process Server und Oracle SOA Suite annotiert werden. Es werden ein direkter Match zu den meisten relationalen Datenbanken und eine semantische Annotation durch OWL- und WSML-Ontologien für einen halbautomatischen Match unterstützt.

Zurzeit beschäftigen wir uns mit der Umsetzung der ganzheitlichen Analyse. Die Matches müssen dazu zwischen Instanzdaten im Audit Trail und dem operativen DWH durchgeführt werden, damit ein Verbund zwischen Prozessdaten und operativen Daten neue und wertvolle Analyseergebnisse liefern kann. Hierfür müssen für die halbautomatische semantische Zusammenführung noch die operativen Datenquellen semantisch annotiert und ein Reasoning Tool bereitgestellt werden.

9 Literaturverzeichnis

- [Ap07] Apache Woden: WSDL2.0 API für Java. Version 1.0, Juli 2007. Verfügbar unter: [¹](http://ws.apache.org/woden)
- [Br05] Bruijn, J. et.al.: Web Service Modeling Language (WSML), W3C Member Submission, 2005. Verfügbar unter: [¹](http://www.w3.org/Submission/WSML/)
- [CS06] Cardoso, J.; Sheth, A.: Semantic Web Services, Processes and Applications. Springer 2006.
- [Ca07] Casati, F. et.al.: A Generic solution for Warehousing Business Process Data. In Proc. of 33rd Int. Conf. on Very Large Data Bases, Österreich 2007.
- [Ch07] Chinnici, R. et.al.: Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language. W3C Recommendation, 2007. Verfügbar unter: [¹](http://www.w3.org/TR/wsd20/)
- [DH05] Doan, A.; Halevy, A.: Semantic Integration Research in the Database Community: A Brief Survey. AI Magazine 2005.
- [FL07] Farrell, J.; Lausen, H.: Semantic Annotations for WSDL and XML Schema. W3C Recommendation, 2007. Verfügbar unter: [¹](http://www.w3.org/TR/sawSDL)
- [He05] Hepp, M. et.al.: Semantic Business Process Management: Using Semantic Web Services for Business Process Management. In Proc. of IEEE Int. Conf. on e-Business Engineering, China 2005.
- [IBM06] IBM: Semantic Tools for Web Services. Verfügbar unter: [¹](http://www.alphaworks.ibm.com/tech/wssem)
- [IBM07] IBM: IBM WebSphere Process Server Version 6.0.2. Verfügbar unter: [¹](http://www-306.ibm.com/software/integration/wps/)
- [ISO06] International Organization for Standardization: ISO Norm 9241-110 – Ergonomics of human-system interaction -- Part 110: Dialogue principles, 2006. Verfügbar unter: [¹](http://www.iso.org/iso/home.htm)
- [Je07] Jena: A Semantic Web Framework for Java. Version 2.5.5, Januar 2007. Verfügbar unter: [¹](http://jena.sourceforge.net)
- [Jo07] Jordan, D. et.al.: Web Service Business Process Execution Language 2.0: Oasis Standard. April 2007. Verfügbar unter: [¹](http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.pdf)

¹ Letzter Zugriff auf aufgeführte Links am 14.05.2008

- [LN06] Leser, U.; Naumann, F.: Informationsintegration – Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen. Dpunkt Verlag, 2006.
- [Ma04] Martin, D. et.al.: OWL-S: Semantic Markup for Web Services. W3C Member Submission, 2004. Verfügbar unter: <http://www.w3.org/Submission/OWL-S/>¹
- [Me05] Meteor-S: Semantic Web Services and Processes. Verfügbar unter: <http://lstdis.cs.uga.edu/projects/meteor-s>¹
- [Or07] Oracle: Oracle SOA Suite Version 10.1.3.1, 2007. Verfügbar unter: <http://www.oracle.com/technologies/soa/soa-suite.html>¹
- [RML08] Radeschütz, S.; Mitschang, B.; Leymann, F.: Matching of Process Data and Operational Data for a Deep Business Analysis. In (Mertins, K. et.al. Hrsg.): Proc. 4th Int. Conf. on Interoperability for Enterprise Software and Applications, Berlin 2008.
- [VD07] Vasudevan, S.; Dave, R.: Oracle Application Server Adapters for Files, FTP, Databases, and Enterprise Messaging User's Guide (10.1.3.1.0), 2007.
- [We05] Weerawarana, S. et.al.: Web Services Platform Architecture: Soap, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging and More. Prentice Hall, New Jersey, USA 2005.
- [Wm98] Workflow Management Coalition: Audit Data Specification. 1998. Verfügbar unter: http://www.wfmc.org/standards/docs/TC-1015_v11_1998.pdf¹
- [Ws06] WSDL4J: WSDL1.1 API für Java. Version 1.6.2, November 2006. Verfügbar unter: <http://sourceforge.net/projects/wsdl4j/>¹
- [Ws07] WSMO4J: WSML API für Java. Version 0.6.1, Juni 2007. Verfügbar unter: <http://wsmo4j.sourceforge.net/>¹
- [Ws08] WSMO Studio: Semantic Web Service Environment for WSMO. Verfügbar unter: <http://www.wsmostudio.org/>¹
- [ZM04] zur Mühlen, M.: Workflow based controlling. Foundation, Design and Application of Workflow-driven Process Information Systems. Logos Verlag, Berlin 2004.

Komponenten- und Serviceorientierung

Leitung und Organisation des Tracks

K. Turowski, Universität Augsburg

Programmkomitee des Tracks

A. Albani, Universität Augsburg/TU Delft
C. Atkinson, Universität Mannheim
A. Bernstein, Universität Zürich
P. Dadam, Universität Ulm
R. Flatscher, Wirtschaftsuniversität Wien
U. Frank, Universität Duisburg-Essen
M. Gaedke, TU Chemnitz
P. Loos, Universität Saarbrücken
V. Nissen, TU Ilmenau
E. Ortner, TU Darmstadt
S. Overhage, Universität Augsburg
W. Pree, Universität Salzburg
C. Rautenstrauch, Universität Magdeburg
E. J. Sinz, Universität Bamberg
C. Szyperski, MS Research/Queensland University
K. Turowski, Universität Augsburg
M. Weske, Hasso-Plattner-Institut
R. Weinreich, Universität Linz

Eine Serviceorientierte Architektur (SOA) für die Integration von RFID-, Sensor- und Lokalisierungsdaten im Facility Management

Daniel Hanhart, Christine Legner, Hubert Österle

Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedbergstrasse 8
CH-9000 St. Gallen
{daniel.hanhart | christine.legner | hubert.oesterle}@unisg.ch

Abstract: RFID-, Sensor- und Lokalisierungstechnologien werden die Abläufe im Facility Management künftig stark verändern. Angesichts der hohen Kosten des Vorort-Einsatzes eines Mitarbeiters und der verteilten Standorte von Immobilien bieten diese Technologien Immobilienbetreibern die Chance, ihre Kostenstrukturen durch hoch integrierte Abläufe zu optimieren und sich im Wettbewerb durch Zusatzleistungen zu differenzieren. Die Integration mobiler und ubiquitärer Technologien in bestehende Systemlandschaften erfordert jedoch die Implementierung zusätzlicher Infrastrukturkomponenten und Schnittstellen. In der Vergangenheit war dies mit dem Einsatz verschiedener, proprietärer Technologien verbunden und führte zu komplexen, schwer beherrschbaren Integrationslösungen. Ausgehend von Anwendungsszenarien in der Instandhaltung von Immobilien zeigt der vorliegende Beitrag, welchen Beitrag Serviceorientierte Architekturen (SOA) für die schnelle und wirtschaftliche Realisierung der vielfältigen, auf RFID-, Sensor- und Lokalisierungstechnologien beruhenden Lösungen leisten können. Dazu konzipieren die Autoren einen serviceorientierten Architekturvorschlag aus fachlicher Sicht und diskutieren den Nutzen im Vergleich zu herkömmlichen Ansätzen.

1 Einleitung

RFID-, Sensor- und Lokalisierungstechnologien stellen für die in einem konsolidierenden Markt operierenden Immobilienbetreiber eine Chance dar, ihre Kostenstrukturen durch hoch integrierte Abläufe zu optimieren und sich im Wettbewerb durch Zusatzleistungen zu differenzieren. Zu den typischen Problemstellungen in der Instandhaltung von Immobilien gehören die geographische Verteilung der Objekte, die hohen Kosten des Vorort-Einsatzes eines Mitarbeiters sowie Medienbrüche in den heutigen Abläufen. Diese Herausforderungen deuten auf beträchtliches Potenzial für mobile und ubiquitäre Lösungen hin [Ea05; Ka05], jedoch erfordert deren Implementierung in der Regel neue Infrastrukturkomponenten und Schnittstellen. In der Vergangenheit wurde dabei jedes Anwendungsszenario mit einer eigenen, proprietären Lösung realisiert. Durch die zu-

nehmende Vielfalt von Szenarien, die RFID-, Sensor- und Lokalisierungstechnologien nutzen, sowie die notwendige Integration mit bestehenden Systemen steigt einerseits die Komplexität der gesamten Systemlandschaft. Andererseits reduziert sich damit deren Beherrschbarkeit und Anpassungsfähigkeit bei neuen Geschäftsanforderungen.

Serviceorientierte Architekturen stellen daher einen viel versprechenden Ansatz für die schnelle und wirtschaftliche Realisierung mobiler und ubiquitärer Anwendungen dar. Ziel dieses Beitrages ist es, die Potenziale einer SOA für die Realisierung der vielfältigen, auf RFID-, Sensor- und Lokalisierungstechnologien basierender Lösungen zu untersuchen. Ausgehend von typischen Anwendungsszenarien im Störfallmanagement bei Immobilien (Abschnitt 2) entwerfen die Autoren einen serviceorientierten Architekturvorschlag (Abschnitt 3). Sie diskutieren dessen Vorteile gegenüber herkömmlichen Ansätzen (Abschnitt 4) und stellen fest, dass SOA eine Voraussetzung für die zukunftsgerichtete Investition in domänenübergreifende mobile und ubiquitäre Lösungen ist. Der Beitrag ist damit dem Forschungsansatz der Design Science [He04] zuzuordnen.

2 Nutzung von RFID-, Sensor- und Lokalisierungsdaten im Störfallmanagement

2.1 Anwendungsszenarien im Störfallmanagement

Aufgabe des Störfallmanagements ist die Annahme, Priorisierung und Verteilung (Disposition) von Störmeldungen sowie die Überwachung der Behebung von Störungen [QG04]. RFID-, Sensor- und Lokalisierungstechnologien finden bereits in folgenden Szenarien Anwendung:

- *Automatische Störmeldung.* Technische Anlagen überwachen ihren Zustand selbst und lösen bei Abweichungen von Sollwerten einen Alarm aus, der zur automatischen Anlage einer Störmeldung im entsprechenden System führt. Stellt das Objekt die Behebung der Störung fest, so meldet es dieses Ereignis und schließt damit die Störmeldung.
- *Mobile Störfallerfassung.* Stellt ein Mitarbeiter eine Störung fest, so erfasst er diese auf seinem mobilen Endgerät und leitet sie zur Behebung weiter. Transponder identifizieren das ausgefallene Objekt. Zusätzlich unterstützt die Lokalisierung des mobilen Endgeräts den Servicemitarbeiter bei der Spezifikation des Schadensortes.
- *Ortsbasierte Personaldisposition.* Die Lokalisierung der Servicemitarbeiter erlaubt dem Disponenten, die verfügbaren Personalressourcen möglichst effektiv einzusetzen und in dringenden Fällen die Aufgabe dem Mitarbeiter zu übergeben, welcher das betreffende Objekt am schnellsten erreichen kann.
- *Mobile Instandhaltungsausführung.* Das mobile Endgerät navigiert den Mitarbeiter in verschiedenen Detaillierungsstufen zum zu reparierenden Objekt. In einem ersten Schritt führt das Gerät den Mitarbeiter auf Basis von Strassen-

karten und groben Lokalisierungsinformation zum Gebäude, in einem zweiten Schritt auf Basis von Gebäudeplänen zur Anlage und in einem letzten Schritt auf Basis von Explosionszeichnungen zum wartenden Objekt bspw. Innerhalb einer komplexen Lüftungs- und Klimaanlage. An die Gebäudekomponenten angebrachte Transponder unterstützen den Servicemitarbeiter bei der eindeutigen Identifikation der Objekte und der Dokumentation der Anwesenheit vor Ort.

2.2 RFID-, Lokalisierungs- und Sensortechnologien

Die Nutzung von RFID-, Lokalisierungs- und Sensortechnologien in den skizzierten Anwendungsszenarien setzt entsprechende Infrastrukturkomponenten sowie Schnittstellen zu bestehenden Anwendungen voraus. Im Einzelnen sind dies:

- *RFID*. Ein RFID-System besteht aus den zwei Komponenten RFID-Transponder und Schreib-/Lesegerät. Ein RFID-Transponder ist ein elektronischer Datenspeicher (Mikrochip), der berührungslos gelesen werden kann. Transponder existieren in unterschiedlichsten Bauformen und werden an dem zu identifizierenden Objekt angebracht. Für den Datenaustausch zwischen Transponder und Lesegerät verwenden RFID-Systeme magnetische oder elektromagnetische Felder.
- *Sensorik*. Gebäude sind für die Steuerung der Temperatur, Beleuchtung, Frischluftversorgung etc. mit einer Vielzahl von Sensoren ausgestattet. Diese sind über definierte Schnittstellen ansprechbar [DH04] und stellen Daten für unterschiedliche Überwachungsfunktionen zur Verfügung [Jo04]. Sie sind über Feldbussysteme vernetzt und kommunizieren über offene industriespezifische Protokolle wie LonTalk¹ oder BACnet², die es gestatten, Geräte unterschiedlicher Hersteller in ein Gebäudeautomationssystem zu integrieren. Auch in RFID-Transpondern können Sensoren integriert sein, z.B. Temperatursensoren, die während des Transports von Waren die Temperatur kontinuierlich aufzeichnen und damit durchgängige Kühlketten dokumentieren.
- *Lokalisierung*. Zur Lokalisierung der Endgeräte und damit ihrer Nutzer bzw. Träger existieren unterschiedliche Verfahren und technische Ansätze. Die Verfahren lassen sich den drei Kategorien *Satellitennavigation*, *netzwerkgestützte Verfahren* und *Trackingverfahren* zuordnen [Ro02]. *Satellitennavigationssysteme* basieren auf der Laufzeitmessung der Funksignale zwischen Endgerät und Satellit. *Netzwerkgestützte Verfahren* nutzen die Information, an welcher Sendestation des Kommunikationsnetzes das Endgerät angemeldet ist. Die

¹ Das Local Operating Network (LON) ist ein Bussystem und wurde 1991 in den USA von der Firma Echelon entwickelt. LonTalk bezeichnet das Kommunikationsprotokoll dieses Bussystems [Bi05].

² BACnet (Building Automation and Control networking) ist ein von der ASHRAE (American Society of Heating, Refrigerating & Air-Conditioning Engineers) entwickeltes Protokoll zur Kommunikation zwischen Infrastruktur- und Applikationsebene in Gebäudeautomationssystemen [Ka04].

Verknüpfung mit der Information über den Standort der Sendestation erlaubt die Positionsbestimmung. *Trackingverfahren* kommen meist innerhalb von Gebäuden zur Anwendung und nutzen Transponder oder Tags zur Lokalisierung. Dabei steht nicht die genaue Position im Vordergrund, sondern die Information, ob ein Objekt oder eine Person (mit Transponder) einen bestimmten Wegpunkt (Lesegerät) passiert haben [Ro02; FD03].

2.3 **Architekturanforderungen**

Tab. 1 zeigt auf, in welchen Anwendungsszenarien im Störfallmanagement die genannten Technologien zum Einsatz kommen. Dies setzt voraus, dass die oben genannten Infrastrukturkomponenten mit den betrieblichen Anwendungssystemen integriert werden. Letztere umfassen typischerweise neben den spezialisieren CAFM (Computer Aided Facility Management)-Systemen ERP-Systeme für Abrechnungsaspekte sowie CAD-Systeme für grafische Gebäudeinformationen.

<i>Arbeitsschritte Technologien</i>	<i>Automatische Störmeldung</i>	<i>Mobile Störfall- erfassung</i>	<i>Ortsbasierte Personal- disposition</i>	<i>Mobile In- standhaltungs- ausführung</i>
<i>RFID</i>		●		●
<i>Sensorik</i>	●			
<i>Lokalisierung</i>		●	●	●

Tabelle 1: Technologien der Anwendungsszenarien

3 **Konzeption einer Serviceorientierten Architektur**

3.1 **Grundlagen und Vorgehen**

Eine SOA erweitert bestehende Ansätze zur Gestaltung von Informationssystemen wie bspw. Client/Server- oder komponentenorientierte Architekturen [Ha06; Si07]. Das im Rahmen dieser Arbeit verwendete Begriffsverständnis einer SOA orientiert sich an den Arbeiten von [W304; Oa05; Er05; He07]: Eine *SOA* ist eine mehrschichtige, verteilte Informationssystem (IS)-Architektur, die Teile der Applikationsarchitektur zur vereinfachten Prozessintegration als Services kapselt und dabei eine Reihe von Designprinzipien berücksichtigt. Letztere umfassen v.a. Schnittstellenorientierung, Autonomie und Modularität, Interoperabilität und Bedarfsorientierung [He07]. Das zentrale Designobjekt einer SOA sind Services, welche die auf Ebene Anwendungssystem vorhandenen Informationsobjekte und Funktionen kapseln und von der technischen und fachlichen Heterogenität der unterschiedlichen Applikationen abstrahieren [Er05].

Der folgende Architekturvorschlag entwickelt das von [He07] konzipierte Architekturmodell für das Facility Management weiter. Dazu prägt er für die konkreten Anwendungsszenarien im Störfallmanagement die Artefakte einer SOA auf den Ebenen An-

wendungssysteme/Domänen, Services, Workflowintegration und Desktopintegration aus. Die SOA-Konzeption wurde in einem systematischen Vorgehen erarbeitet und in Expertengesprächen validiert: Aus den Anwendungsszenarien aus Abschnitt 2.1 leiten sich die Anforderungen an mobile und ubiquitäre Services ab. Die Szenarien wurden daher auf den Ebenen Desktopintegration (als Taskflows) und Workflowintegration (als applikationsübergreifender Workflow) modelliert und einer Liste von Servicekandidaten hergeleitet. Basis bildet die Anwendungssysteme bzw. Domänen, die als Serviceanbieter auftreten. Die Ableitung der Services erfolgte unter der Annahme, dass der externe Zugriff auf eine Domäne ausschließlich über Serviceschnittellen zu implementieren ist, während eine domänenintern bereitgestellte IT-Unterstützung in der Regel in der gleichen Applikation realisiert ist [He07].

3.2 Ebene Anwendungssystem

Die Ebene Anwendungssystem stellt die Ressourcensicht dar, welche die in Applikationen vorhandene und zur Unterstützung betrieblicher Aktivitäten nutzbaren fachlichen Funktionen und Informationsobjekte umfasst. Um von den Spezifika und der Vielfalt der möglichen Realisierungsvarianten zu abstrahieren, gruppieren wir fachlich zusammengehörende Funktionen und Informationsobjekte in Domänen, die als Serviceanbieter auftreten. Für das Facility Management wurden durch entsprechende Gruppierung von Geschäftsobjekten und Funktionalitäten folgende Domänen abgeleitet (für eine ausführliche Herleitung sei auf [Ha07] verwiesen):

- Die Domäne *Gebäudedaten (GD)* bildet das Gebäudemodell mit den Gebäudekomponenten und den Raumstrukturelementen ab. Sie umfasst alle wichtigen alphanumerischen und grafischen- bzw. CAD-Daten zur Beschreibung von Gebäuden.
- Die Domäne *Kundendaten (KD)* enthält sämtliche Geschäftsobjekte, die Kunden und die dazugehörigen Verträge betreffen.
- Die zentralen Geschäftsobjekte der Domäne *Analyse & Optimierung (AO)* sind alle Zustands- und Nutzungsdaten, die der Prozess Störfallmanagement verwendet.
- Die Domäne *Auftragsabwicklung (AA)* umfasst die zur Planung, Disposition, Abwicklung und Rückmeldung von Tätigkeiten bzw. Aufträgen notwendigen Geschäftsobjekte und Funktionen. Dies sind insbesondere die Geschäftsobjekte Instandhaltungsplan, Instandhaltungsauftrag, Leistungsverzeichnis, Reinigungsauftrag, Rundgang und Umzugsauftrag.
- Die Domäne *Materialwirtschaft (MM)* fasst die zentralen Elemente für die Lagerbewirtschaftung und den Materialeinkauf zusammen. Dies sind die Geschäftsobjekte Material, Lieferant, Reservation, Bestellung, (Lieferanten)-Rechnung und Lagerplatz.
- Die Domäne *Störfallmanagement (SFM)* enthält schliesslich die Geschäftsobjekte Störmeldung und Schadenskatalog sowie die für Fehleranalyse wichtigen Wissensdatenbanken bestehend aus Problem-/Lösungspaaren.

3.3 Ebene Desktopintegration

Die Ebene Desktopintegration unterstützt den Benutzer bei der Durchführung einzelner Arbeitsschritte (Tasks). Dabei bilden Arbeitsschritte atomare Einheiten, die wiederum zu Aktivitäten zusammengefasst werden. Services unterstützen die manuelle Arbeitsschritte dadurch, dass sie dem Benutzer alle für die Erledigung notwendigen Daten und Funktionen integriert zur Verfügung stellen.

Portale sind heute die häufigste Realisierungsform der Desktopintegration. Sie eignen sich im Facility Management insbesondere für die Unterstützung des Vor-Ort-Einsatzes von Mitarbeitern. RFID-, Sensor- und Lokalisierungstechnologien unterstützen bei den Anwendungsszenarien aus Abschnitt 2.1 folgende manuellen Arbeitsschritte: (1) „Objekt identifizieren“ bei der mobilen Störfallerfassung bzw. mobilen Instandhaltungsausführung; (2) „Mitarbeiter lokalisieren“ bei der mobilen Störfallerfassung zur Spezifikation des Schadensortes, bei der ortsbasierten Personaldisposition zur Bestimmung des geeigneten Mitarbeiters für einen dringenden Störfall und bei der mobilen Instandhaltungsausführung für die Navigation zum Objekt. Tab. 2 ordnet den identifizierten Arbeitsschritten die benötigten Services und den entsprechenden Domänen zu. Der Arbeitsschritt „Objekt identifizieren“ wird durch ein im mobilen Endgerät integriertes RFID-Lesegerät unterstützt, das die Daten des RFID-Transponders ausliest und bestimmte Felder der Anwendung auf dem mobilen Endgerät füllt. Der Serviceaufruf kann somit geräteintern abgebildet werden. Den Service zur Ermittlung der aktuellen Position kann ein mit einem GPS-Empfänger ausgestattetes Endgerät ebenfalls intern erbringen. Basiert die Lokalisierung auf der Zellinformation in Mobilfunknetzen, so liefert der Mobilfunkanbieter oder bei Trackingverfahren ein RFID-System die zur Endgeräteidentifikation aktuelle Koordinate zurück. In diesen Fällen sind die Lokalisierungsdaten über einen Service der Domäne Analyse & Optimierung (AO) verfügbar.

Nr.	Arbeitsschritt	Service	Eingabe	Ausgabe	Domäne
1	Objekt identifizieren	-	-	[GK-ID]	geräteintern
2	Mitarbeiter lokalisieren	Position ermitteln	[Endgerät-ID]	Koordinate	geräteintern/AO
Legende: ID Schlüsselattribut eines Geschäftsobjekts, GK Gebäudekomponente [] Liste von Geschäftsobjekten, AO-Analyse & Optimierung					

Tabelle 2: Zuordnung Services zu Arbeitsschritten

3.4 Ebene Workflowintegration

Die Ebene Workflowintegration bildet Prozessabläufe als Workflows ab und trennt so die Prozessablaufsteuerung (Kontrollfluss) von der Applikationslogik [Oe95; Sc02]. Ein Workflow kann sich sowohl aus manuell von Personen als auch automatisiert von Applikationen ausgeführten Aktivitäten zusammensetzen [Wf99; Vo03].

Die für das Störfallmanagement entworfenen Workflows basieren auf dem von [BN03] vorgeschlagenen Referenzprozess und realisieren durchgängige, elektronisch integrierte Prozesse. Die Definition geeigneter Services soll eine lose Kopplung der Domänen und eine flexible Integration der Lösungen auf Basis von RFID-, Sensor- und Lokalisierungstechnologien erlauben. Die folgende Darstellung fokussiert daher auf die domänen-

übergreifenden Prozessbestandteile und die mit RFID-, Sensor- und Lokalisierungstechnologien unterstützten Aktivitäten und erzeugten Ereignisse.

Das Erkennen einer Störung erfolgt entweder manuell durch eigene Mitarbeiter bzw. Kunden oder automatisiert durch Gebäudekomponenten und löst den Störfallmanagementprozess aus. Diese unterschiedlichen Startereignisse haben Konsequenzen auf die Servicekonzeption und sind daher als Prozessvarianten automatisierte bzw. manuelle Störmeldung dargestellt (vgl. Abbildung 2).

Diese Varianten unterscheiden sich auch im Abschluss des Prozesses. Die von Gebäudekomponenten angelegten Störmeldungen sind automatisch, d.h. vom entsprechenden Ereignis derselben Gebäudekomponente, zu schließen. Wird auf diese Forderung verzichtet, kann dies zu inkonsistenten Systemzuständen führen. Schließt bspw. ein Mitarbeiter die Störmeldung zu einer Gebäudekomponente trotz eines weiterhin bestehenden Alarms des Gebäudeautomationssystems, so wird dieser Alarm nicht weiter behandelt. Für die Bearbeitung einer Störmeldung sind die zu ergreifenden Maßnahmen und die verantwortliche Organisationseinheit festzulegen. Je nach betroffener Komponente kann eine Entstörung per Fernwartung oder durch einen Vor-Ort-Einsatz erfolgen. Ist ein Vor-Ort-Einsatz notwendig, so ist ein entsprechender Instandhaltungsauftrag anzulegen und zu disponieren. Anschließend muss der Mitarbeiter den Auftrag abarbeiten und zurückmelden. Beheben die getroffenen Maßnahmen die Störung nicht, so leitet die für die Störmeldung verantwortliche Organisationseinheit die Meldung weiter, um ggf. einen erneuten Entstörungsversuch einzuleiten. Ist die Störung behoben, so kann die Störmeldung entweder durch die betroffene Gebäudekomponente, durch Schließen des zugehörigen Instandhaltungsauftrags oder manuell geschlossen werden.

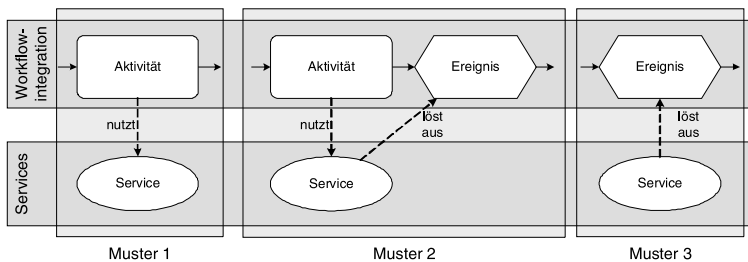


Abbildung 1: Interaktionsmuster zwischen Services, Aktivitäten und Ereignissen

Serviceinteraktionsstile

Ein Service kann Ereignisse auslösen und/oder wird von einzelnen Aktivitäten genutzt. Der Interaktionsstil eines Service beeinflusst die Anzahl der Serviceinteraktionen für Sequenzen von Aktivitäten und Ereignissen. [He07] unterscheidet die Interaktionsstile synchron oder asynchron sowie Notification oder Request/Response. Aus den unterschiedlichen Interaktionsstilen ergeben sich die drei in Abb. 1 dargestellten Interaktionsmuster zwischen Ereignis-/Aktivitätssequenzen und Services. Aktivitäten können sowohl synchrone wie auch asynchrone Notification- (Muster 1) oder Request/Response-Services (Muster 2) nutzen. Auslöser für Ereignisse können in Analogie dazu sowohl synchrone wie auch asynchrone Request/Response-Services (Muster 2) sowie Notification-Services (Muster 3) sein.

Der passende Interaktionsstil für einen Service ist vom übertragenen Datenvolumen, Anforderungen an Antwortzeiten und weiteren Faktoren abhängig. Er ist damit abhängig von der konkreten Anwendung des Services festzulegen. Der vorliegende Architekturvorschlag unterscheidet nur zwischen Request/Response- und Notification-Services, verzichtet aber auf die Unterscheidung zwischen synchroner und asynchroner Kommunikation, da diese Unterscheidung die fachliche Konzeption eines Service nicht beeinflusst (s. Abb. 1). Weiter ist eine Sequenz von Aktivität und Ereignis immer mit einem Request/Response-Service (Muster 2) modelliert. Bei Bedarf ist im konkreten Anwendungsfall ein Request/Response-Service (Muster 2) problemlos in eine Sequenz von zwei Notification-Services (Muster 1 und 3) zu übersetzen. Schließlich verzichtet die hier vorliegende fachliche Modellierung auf das Abbilden von Fehlerfällen (Serviceaufruf bricht ab, zu einem Schlüssel existiert kein Eintrag im System etc.).

Trennung von Daten- und Kontrollfluss in Workflows

Der Informationsfluss entlang eines Workflows unterteilt sich in Kontrollfluss sowie den eigentlichen Datenfluss. Die Kontrolldaten sind die zentralen Informationen der Ebene Workflowintegration und deshalb auch von ihr zu verwalten, während die Daten zu Geschäftsobjekten in den Anwendungssystemen der Domänen enthalten sind. Das Anlegen einer Prozessinstanz, die sowohl Kontroll- wie auch Geschäftsdaten eines Prozessdurchlaufs enthält, würde zu einer redundanten Datenhaltung führen und ist zu vermeiden. Die Steuerungskomponente auf Ebene Workflowintegration soll daher nur die *Schlüsselattribute* der Geschäftsobjekte führen, während die restlichen Geschäftsdaten in den Domänen verbleiben und dort mit den jeweiligen Applikationen bearbeitet werden [Vo03]³. Die vorgeschlagenen Workflows modellieren den Datenfluss als Austausch von Schlüsselattributen zwischen den Aktivitäten und Ereignissen (s. Abb. 2). Als Konsequenz sind die domänenübergreifenden Datenflüsse auf Ebene Service abzubilden.

³ Im Gegensatz dazu lassen andere Autoren die Abbildung eines über die Schlüsselattribute hinausgehenden Datenflusses auf Ebene Workflow zu [Zu04]. Weiter vereinfacht die strikte Trennung von Kontroll- und Datenfluss die Analyse und verbessert das Verständnis von Workflows [KHA03].

	Beschreibung	Typ	Eingabe	Ausgabe	Geschäftsobjekte	Domäne ⁴
1	Störung automatisch erkannt	N		GK-ID	Zustandsdaten	AO
2	Störmeldung generieren	RR	GK-ID	Meldungs-ID	Störmeldung Zustandsdaten	CO 3 & 4
3	Zustandsdaten lesen	RR	GK-ID	Zustandsdaten	Zustandsdaten	AO
4	Störmeldung anlegen	RR	Störmeldung	Meldungs-ID	Störmeldung	SFM
5	Fernwartung ist durchzuführen	N		Meldungs-ID	Störmeldung	SFM
6	Vor-Ort-Einsatz ist notwendig	N		Meldungs-ID	Störmeldung	SFM
7	Instandhaltungsauftrag zu Störmeldung anlegen	RR	Meldungs-ID	Auftrags-ID	IH-Auftrag Störmeldung	CO 8 & 9
8	Störmeldung lesen	RR	Meldungs-ID	Störmeldung	Störmeldung	SFM
9	Instandhaltungsauftrag anlegen	RR	IH-Auftrag	Auftrags-ID	IH-Auftrag	AA
10	Störung ist nicht behoben	N		Meldungs-ID	Störmeldung	SFM
11	Störung zu Objekt ist behoben	N		GK-ID	Zustandsdaten	AO
12	Störmeldung ändern	RR	Störmeldung	Meldungs-ID	Störmeldung	SFM
13	IH-Auftrag abgeschlossen	N		Auftrags-ID	IH-Auftrag	AA
14	Störmeldung manuell geschlossen	N		Meldungs-ID	Störmeldung	SFM
Legende: N Notification; RR Request/Response ID Schlüsselattribut eines Geschäftsobjekts GK Gebäudekomponente, RS Raumstrukturelement, IH Instandhaltung, () optionale Inputparameter, [] Liste von Geschäftsobjekten GD-Gebäudedaten, KD-Kundendaten, AA-Auftragsabwicklung, AO-Analyse & Optimierung, SFM-Störfallmanagement, MM-Materialwirtschaft, CO-Composite						

Tabelle 3: Services zur Abbildung des Workflow ungeplante Instandhaltung

Ausnahmen zur 1:1-Realisierung von Aktivitäten und Ereignissen sind:

- Der Service *Störmeldung generieren* (2) soll die aktuellen Zustandsdaten der Gebäudekomponente auslesen und diese Information in die Störmeldung übertragen (z.B. detaillierte Fehlermeldung oder Logdaten des ausgefallenen Objektes). Da die Ebene Workflowintegration nur die Identifikationsnummer der Gebäudekomponente überträgt, ist der Datenfluss als Abfolge von zwei Serviceaufrufen abzubilden. Der erste Service liest den aktuellen Zustand der betroffenen Gebäudekomponente aus (3), der dann Basis für die Anlage der Störmeldung durch den zweiten Service ist (4). Wie in Abbildung 4 dargestellt, koordiniert der Service *Störmeldung generieren* (2) diese zwei Serviceaufrufe, erwartet als Eingabeparameter die Identifikationsnummer der defekten Gebäudekomponente und liefert die Identifikation der angelegten Störmeldung zurück.

⁴ Bei komponierten Services findet sich jeweils eine Referenz auf die Services, aus welchen er sich zusammensetzt.

- Analog liest der bereits in Abbildung 3 erwähnte Service *Instandhaltungsauftrag zu Störmeldung anlegen* (7) zuerst die Störmeldung aus (8) und legt anschließend den Instandhaltungsauftrag mit Informationen aus der Störmeldung an (9).
- Das Ereignis *Störung zu Objekt ist behoben* der Prozessvariante automatisierte Störfallanlage (s. Abb. 4) ist mit einem Notification-Service (11) implementiert, der als Ausgabeparameter die Identifikationsnummer der erneut betriebsbereiten Gebäudekomponente hat. Anschließend setzt die Aktivität „Störmeldung schließen“ mit dem Service *Störmeldung ändern* (12) den Status der Störmeldung auf abgeschlossen. Dieser Service benötigt als Eingabe die Störmeldungsidentifikation. Diese Meldungs-ID liefert nicht der auslösende Notification-Service (11), sondern ist auf Ebene Workflowintegration bei der Prozessinstanz hinterlegt. Die Prozessinstanz ist über die Identifikation der Gebäudekomponente, welche der auslösende Notification-Service (11) liefert, identifizierbar.

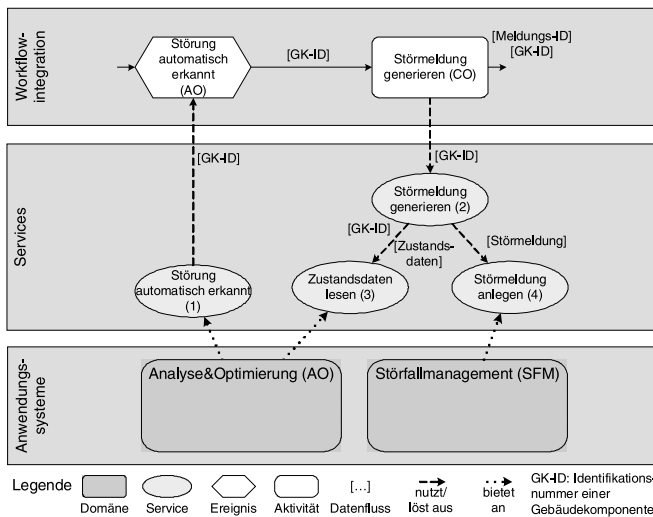


Abbildung 4: Serviceinteraktion für automatisch generierte Störmeldungen

- Bei der Prozessvariante manuelle Störmeldung ist analog zu obigen Erläuterungen die Prozessinstanz über die Auftrags-ID des Notification-Services *Instandhaltungsauftrag ist abgeschlossen* (13) identifizierbar. Für die Prozessinstanz ist auf Ebene Workflowintegration die Meldungs-ID hinterlegt, und die Aktivität „Störmeldung schließen“ kann mit demselben Service *Störmeldung ändern* (12) realisiert werden.
- Der Notification-Service *Störmeldung manuell geschlossen* (14) ist notwendig, damit die Prozessvariante manuelle Störfallanlage immer im Endzustand „Störmeldung geschlossen“ terminiert.

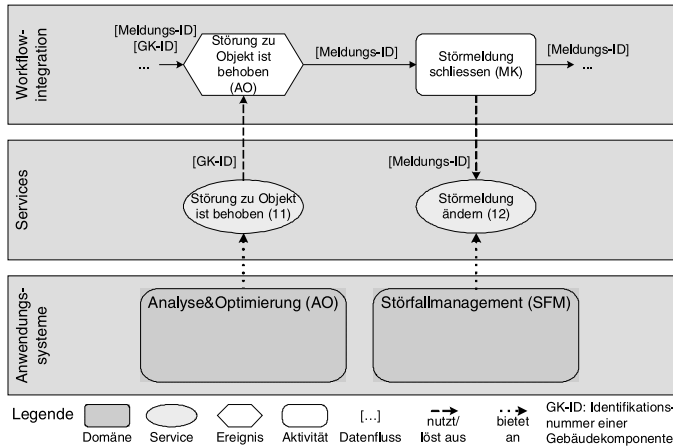


Abbildung 5: Serviceinteraktion für das automatische Schließen von Störmeldungen

4 Zusammenfassung und Bewertung

Der vorliegende Beitrag konzipierte die fachlogische Sicht einer SOA für typische Anwendungsszenarien von RFID-, Sensor- und Lokalisierungstechnologien im Facility Management. Dazu wurden Taskflows und Workflows für das Szenario Störfallmanagement modelliert und insgesamt 15 Servicekandidaten abgeleitet, mit denen sich die wichtigsten mobilen und ubiquitären Szenarien realisieren lassen. Bewusst verzichtet der serviceorientierte Architekturvorschlag auf ein komplettes Redesign der IT-Unterstützung im Facility Management, sondern konzentriert sich auf die Verwendung von Services vielmehr zur Prozessintegration an den Domänengrenzen. Vergleicht man den serviceorientierten Architekturvorschlag mit der herkömmlichen Realisierung von Lösungen auf Basis von RFID-, Sensor- und Lokalisierungstechnologien über proprietäre Schnittstellen, so entstehen Vorteile in drei Bereichen:

- *Wiederverwendung.* Mit einer hohen Wiederverwendung bestehender Schnittstellenfunktionalität ist insbesondere bei Notification-Services zu rechnen, wenn die erzeugten Ereignisse mehrere Aktivitäten auslösen, also unterschiedliche Prozesse steuern. Durch die Trennung von Ereignissen und Aktivitäten besteht die Möglichkeit, Ereignisse und damit zugehörige Notification-Services mehrfach zu verwenden. Beispielsweise kann eine automatisierte Störmeldung einerseits die Störungsbehebung wie auch eine Reinigung auslösen. Umgekehrt können manuelle Aktivitäten die von einer automatisierten Aktivität genutzten Services wieder verwenden. So ist bspw. der Service *Störmeldung anlegen* in den beiden Varianten automatische Störmeldung und mobile Störfallerfassung nutzbar.
- *Entkopplung.* Der serviceorientierte Architekturvorschlag reduziert Abhängigkeiten zwischen Applikationen und erhöht so die Beherrschbarkeit der Systemlandschaft. Im Anwendungsfall Störfallmanagement macht sich diese Entkopplung

lung u.a. durch die Trennung von Ereignissen und Aktivitäten bemerkbar, die eine Integration neuer Lösungen auf Basis von RFID, Sensorik und Lokalisierung vereinfacht. Eine neu umzusetzende Lösung muss einzig einen Notification-Service anbieten, welcher auf Ebene Workflowintegration mit den auszulösenden Aktivitäten verknüpft wird. Die Entkoppelung der Domänen vereinfacht spätere Anpassungen der Anwendung bspw. bei Technologiewechseln. Solange die ausgelösten Ereignisse stabil sind, beschränken sich die Anpassungen auf die Neuimplementierung der Notification-Services und damit auf eine Domäne. Im Facility Management ist insbesondere die Entkopplung der von der Domäne Analyse & Optimierung erzeugten Ereignisse von den Domänen Störfallmanagement oder Auftragsabwicklung von Vorteil.

- *Realisierung domänenübergreifender Anwendungen.* Sowohl die technische Standardisierung wie auch die konzeptionelle Vereinheitlichung beim Service-design erlauben, Funktionen unterschiedlicher Domänen zur Realisierung einer Applikation so zu verwenden, als ob diese von einer einzigen Anwendung zur Verfügung gestellt würden. Beispiele solcher Applikationen sind Portale oder mobile Anwendungen, welche auf Ebene Desktopintegration alle zur Erledigung einer Aktivität notwendigen Funktionen zusammenführen. So kann bspw. die Anwendung für das Anwendungsszenario „mobile Störfallerfassung“ sowohl auf Lokalisierungs- und Zustands-/Nutzungsdaten der Domäne Analyse & Optimierung zugreifen, wie auch in der Domäne Störfallmanagement eine Störmeldung anlegen. Auch Workflowmanagementsysteme können als Service gekapselte Funktionalität unterschiedlicher Domänen nutzen. Mit der Zuweisung manueller Aktivitäten zu Personen und der Verwaltung von Aktivitäten unterschiedlicher Domänen ist die Realisierung eines zentralen domänenübergreifenden Arbeitsvorrates möglich.

Angesichts der immer integrierteren Abläufe im Facility Management kommt den hier angesprochenen Nutzenpotenzialen besondere Bedeutung zu, da sich aktuelle Systemlandschaften durch eine steigende Anzahl von Systemen mit einer exponentiell steigenden Menge von Integrationsbeziehungen auszeichnen. Eine Serviceorientierte Architektur kann somit die Voraussetzung für eine zukunftsgerichtete Investition in Lösungen auf Basis von RFID-, Sensor und Lokalisierungstechnologien angesehen werden.

5 Literaturverzeichnis

- [BN03] Becker, J., Neumann, S.: Referenzmodell für Workflow-Applikationen in technischen Dienstleistungen, in: Bullinger, H.-J., Scheer, A.-W. (Hrsg.), Service Engineering: Entwicklung und Gestaltung innovativer Dienstleistungen, Springer, Berlin etc. 2003, S. 619-645.
- [Bi05] Biehlig, C.: Haustechnik für Verwalter, Vermieter und Makler, Haufe, München 2005.
- [DH04] Diekmann, T.: Hagenhoff, S., Eignung von Web Services für Embedded Devices in betrieblichen Anwendungssystemen, Working Paper, Universität Göttingen, Göttingen 2004.

- [Ea05] Eastwood, G.: The Enterprise Mobility Market Outlook – Fast growth in mobile solutions, devices, email and security, Business Insights, 2005.
- [Er05] Erl, T.: Service-Oriented Architecture, Prentice Hall, New York 2005.
- [FD03] Fleisch, E., Dierkes, M.: Ubiquitous Computing aus betriebswirtschaftlicher Sicht, in: Wirtschaftsinformatik, 45, 2003, Nr. 6, S. 611-620.
- [Ka05] Habermann, K.: AVENTON Mobile Business Assistant – Creating and Optimizing „Mobile Business Processes“, in: Wirtschaftsinformatik, 47, 2005, Nr. 1, S. 55-62.
- [Ha07] Hanhart, D.: Mobile und Ubiquitous Computing – Anwendungen im Facility Management und serviceorientierter Architekturvorschlag, Dissertation, Universität St. Gallen, St. Gallen 2007.
- [Ha06] Hasselbring, W.: Software-Architektur, in: Informatik Spektrum, 29, 2006, Nr. 1.
- [He04] Hevner, A. R. et al.: Design Science in Information Systems Research, in: MIS Quarterly, 28, 2004, Nr. 1, S. 75-105.
- [He07] Heutschi, R.: Serviceorientierte Architektur – Architekturstil und Ansätze für die Umsetzung, Dissertation, Universität St. Gallen, St. Gallen 2007.
- [Jo04] Jones, T. M.: TCP/IP Application Layer Protocol for Embedded Systems, Laxmi, New Dehli 2004.
- [Ka04] Karbach, A.: Gebäudeleittechnik und technisches Gebäudemanagement, in: Regelungstechnik, A.d.P.f. (Hrsg.), Digitale Gebäudeautomation, Springer, Berlin 2004, S. 315-371.
- [KHA03] Kiepuszewski, B., Hofstede, A. H. M. t., van der Aalst, W. M. P.: Fundamentals of control flow in workflows, in: Acta Informatica, 39, 2003, Nr. 3, S. 143-209
- [Oe95] Österle, H.: Business Engineering: Prozess- und Systementwicklung, Band 1: Entwurfstechniken, 2. Auflage, Springer, Berlin 1995.
- [QG04] Quadt, M., Görze, R.: Geschäftsprozesse im Facility Management und ihre Abbildung in der IT, in: May, M. (Hrsg.), IT im Facility Management erfolgreich einsetzen: Das CAFM-Handbuch, Springer, Berlin 2004, S. 37-79.
- [Ro02] Roth, J.: Mobile Computing: Grundlagen, Technik, Konzepte, dpunkt-Verlag, Heidelberg 2002.
- [Sc02] Scheer, A.-W.: ARIS – vom Geschäftsprozess zum Anwendungssystem, Springer, Berlin 2002.
- [Si07] Siedersleben, J.: SOA revisited: Komponentenorientierung bei Systemlandschaften, in: Wirtschaftsinformatik, 49, 2007, Nr., S. 110-117.
- [Vo03] Vogler, P.: Prozess- und Systemintegration: Umsetzung des organisatorischen Wandels in Prozessen und Informationssystemen, Habilitationsschrift, Habilitation, Universität St. Gallen, 2003.
- [Wf99] WfMC: Terminology & Glossary, http://www.wfmc.org/standards/docs/TC-1011_term_glossary_v3.pdf, 08.06.2006.
- [Zu04] Zur Muehlen, M.: Workflow-based Process Controlling. Foundation, Design, and Application of workflow-driven Process Information Systems., Logos, Berlin 2004.

Evaluation von Konzepten zur Integration von UDDI-Registries unter Berücksichtigung von Marktaspekten

Nico Brehm, Liane Haak, Jorge Marx Gómez

Abteilung Wirtschaftsinformatik I/Very Large Business Applications
Department für Informatik
Carl von Ossietzky Universität Oldenburg
Ammerländer Heerstraße 114-118
26129 Oldenburg
{brehm | haak | marx-gomez}@wi-ol.de

Abstract: Bei der Verwendung eines zentralen Verzeichnisdienstes als Bestandteil einer Serviceorientierten Architektur entstehen Probleme, deren Lösung bei der Betrachtung von Web Services als vermarktbar Softwarekomponenten besondere Wichtigkeit zukommt. Verzeichnisdienste, wie beispielsweise UDDI-Registries, sind in diesem Zusammenhang als Marktunterstützungsdienste zu betrachten, die sowohl Anbietern als auch Nutzern von Web Services nicht nur Mehrwerte bieten, sondern essentielle Grundlage für die Etablierung von Web Service-Märkten sind. Diese Erkenntnis erfordert die Entwicklung und Bewertung von Konzepten, die den Betrieb von UDDI-Registries als Aufgabe der Teilnehmer eines Marktes selbst betrachten. Im Beitrag werden drei komplementäre Konzepte zum Betrieb von integrierten UDDI-Registries gegenübergestellt und bezüglich ihrer Vor- und Nachteile im Kontext eines offenen Web Service-Marktes bewertet.

1 Einleitung

Im Unterschied zu Softwarekomponenten, deren Integration sich auf die „lokale“ Installation innerhalb eines übergeordneten Softwaresystems bezieht, können Web Services¹ als elektronische Dienstleistungen betrachtet werden [BS03], die durch externe Systeme erbracht werden [BMZ08].

¹ Ein Web Service ist aus der Sicht eines Nutzers eine Softwarekomponente, die ausschließlich Dienste anbietet (exportiert) und keine Dienste nachfragt (importiert). Die Schnittstellen der angebotenen Dienste werden durch Konstrukte standardisierter XML-Sprachen spezifiziert. Die Nutzung eines Dienstes ist abhängig von einem Austausch von Nachrichten, deren Struktur durch die Schnittstellenspezifikation festgelegt wird. Nachrichten werden grundsätzlich über ein bestehendes Netzwerk übertragen. Aus der Sicht eines Anbieters ist ein Web Service ein Anwendungssystem, dass die benannten Eigenschaften aus Nutzersicht für jeden Nutzer erfüllt. Die Verfügbarkeit aller importierten Dienste wird durch den Anbieter in Eigenverantwortung (autonom) organisiert [BMZ08].

Web Service-Nutzer und Web Service-Anbieter befinden sich dabei in einer Anbieter-Nachfrager-Beziehung, die mit der Beziehung zwischen den Akteuren in konventionellen Märkten vergleichbar ist. Eine bedeutende Voraussetzung für den Erfolg der komponentenorientierten Anwendungsentwicklung im Allgemeinen ist das Vorhandensein eines Komponentenmarktes ([Ov04]; [Te03]; [HW02]; [JT03]; [Ha06]).

Unter der Verallgemeinerung verschiedener Modelle zur dieser Thematik kann ein Komponentenmarkt zunächst als Ort (Marktplatz) des Aufeinandertreffens zwischen einem Angebot von Komponenten und der Nachfrage nach Komponenten definiert werden [Ha06]. In Anlehnung an die Eigenschaften von Märkten im Allgemeinen ist die Standardisierung und in diesem Zusammenhang der genauen Spezifikation von Komponenten, als die auf dem Markt gehandelten Güter, besonders wichtig [Ha06]. Verschiedene Anbieter von Komponenten können in Konkurrenz zueinander auftreten und die angebotenen Komponenten können durch die Verfügbarkeit von Informationen sowohl hinsichtlich ihrer Anforderungserfüllung (aus der Sicht der Nachfrager) bewertet [Ha06], als auch miteinander verglichen werden.

Aufbauend auf den Ausführungen von Merz [Me99] erscheint eine Nachbildung eines Marktkonstrukts, d.h. ökonomischer Beziehungen zwischen den Akteuren, auch für Web Services realitätsnah, da diese als elektronische Dienste aufgefasst werden können, deren Nutzung kostenpflichtig ist². Nach Merz schließt ein elektronischer Markt sämtliche Softwarekomponenten ein, die das Zustandekommen einer Handelstransaktion unterstützen [Me99]. Hierbei sind Anwendungssysteme, die Kommunikationsinfrastruktur als auch weitere unterstützende Systeme [Me99], wie etwa zur Abrechnung oder zur Speicherung von Daten über den Markt und seine Teilnehmer zu nennen.

In Anlehnung an die Darstellung der Architektur eines elektronischen Marktsystems nach [ZK95] und den Ausführungen in [Me99] stellt sich eine Architektur für ein Web Service-Marktsystem wie in Abbildung 1 dar. Ein Marktsystem, dass den Handel mit Web Services ermöglicht, besteht dabei zunächst aus einer Zugangssoftware für die Nutzer von Web Services als Teil eines übergeordneten Anwendungssystems, sowie einer Zugangssoftware für die Anbieter von Web Services. Weiterhin sind Komponenten zur Kommunikationsunterstützung notwendig, die wiederum auf einem bestehenden Transportmedium aufsetzen.

² Grundsätzlich beschreiben die Einzelschritte 1) Angebot, 2) Annahme, 3) Leistung und 4) Zahlung einen Kauf [Me99], dessen zugrunde liegender Vertrag (Kaufvertrag) in die beiden Willenserklärungen des Anbieters, durch die Bereitstellung eines Dienstes, und des Nutzers, durch den Aufruf eines Dienstes, aufgeteilt werden kann. Die Spezifikation eines Web Services beinhaltet darauf aufbauend den Vertragsinhalt, durch den sich der Anbieter auf der einen Seite selbst zur Erbringung einer Dienstleistung verpflichtet. Weiterhin wird ein Nutzer durch die Spezifikation eines Web Services über den zu zahlenden Nutzungspreis informiert. Es gilt dabei zu beachten, dass der Vertragsinhalt nicht für jede Anbieter-Nutzer-Beziehung gleich sein muss, so dass die Spezifikation des Vertragsinhalts auch erst auf Anfrage des potentiellen Nutzers möglich ist.

Zusätzlich sind weitere Unterstützungsdienste bezogen auf das Zustandekommen und die Durchführung von Markttransaktionen notwendig. Beispielhaft seien hier Komponenten zur Suche nach Web Services hervorgehoben, die in den weiterführenden Abschnitten vertiefend betrachtet werden sollen.

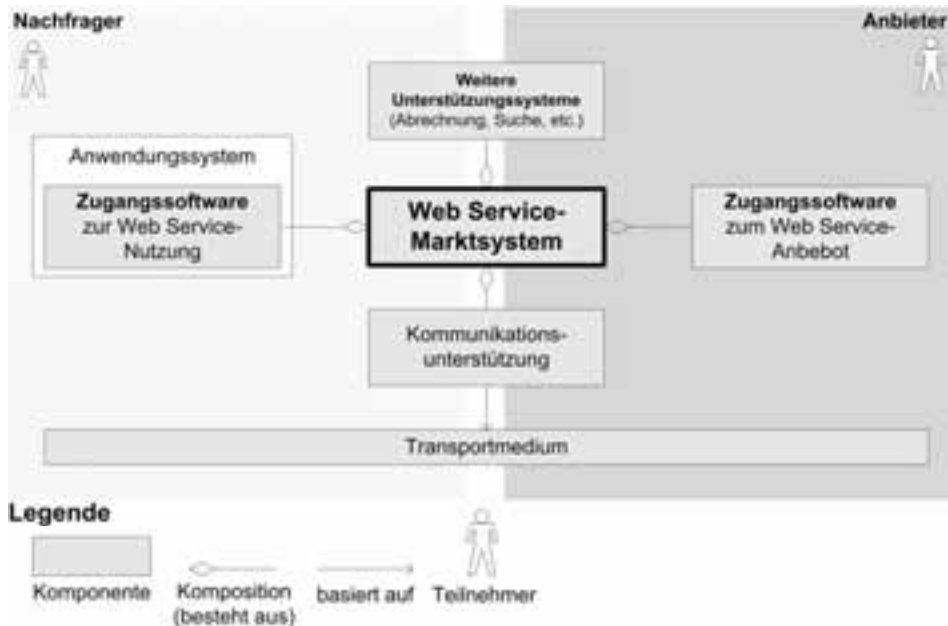


Abbildung 1: Architektur eines Web Service-Marktsystems

Bezugnehmend auf die dargestellten Eigenschaften eines Web Service-Marktsystems sind Suchdienste für Web Services auch als essentielle Bausteine zur Umsetzung Web Service-basierter SOA zu betrachten, deren Aufgabe es ist, Web Services in einem Netzwerk bekannt zu machen bzw. zu publizieren. Der hierbei am häufigsten diskutierte Standard ist UDDI (Universal Description, Discovery and Integration) und bietet Schnittstellen zur Hinterlegung von Web Service-Zugriffspunkten und zur Suche nach benötigten Web Services. Über diese Schnittstelle werden die Web Services vom Dienstanbieter in eine UDDI-Registry eingetragen und können über standardisierte Suchanfragen von Clients (Dienstnutzer) aufgefunden werden. Die UDDI-Registry fungiert dabei in konventionellen Ansätzen in der Regel als zentraler Verzeichnisdienst. Die dienstbezogenen Informationen, die für die spätere Suche nach Web Services veröffentlicht werden können, sind in Bezug auf ihre inhaltliche Ausrichtung verschiedenen Bereichen zugeordnet, die auch als *White Pages*³, *Yellow Pages*⁴ und *Green Pages*⁵ bezeichnet werden [OT05].

³ Hierdurch werden allgemeine Informationen über den Anbieter von Web Services, wie z.B. Ansprechpartner oder Adressdaten beschrieben.

Neben diesen konventionellen Informationen über Web Services existieren weiterhin Vorschläge zur Veröffentlichung erweiterter Informationen. Overage et al. schlagen beispielsweise die Berücksichtigung von *Blue Pages* zur Spezifikation der Funktionalität und *Grey Pages* zur Spezifikation der Qualität von Web Services vor [OT05].

2 Problemstellung

Der zentralisierte Charakter der Infrastruktur von Web Service-basierten Anwendungssystemen ergibt sich durch den konventionellen Ansatz, eine einzige Instanz eines Verzeichnisdienstes zu berücksichtigen. Der Betrieb einer UDDI-Registry ist dadurch abhängig von einer Verwaltungsstelle, die aus organisatorischer Sicht eine Zentrale darstellt.

Die zentralisierte Verwaltung von Web Service-Angeboten hat in diesem Kontext zunächst folgende Vorteile:

- *Gemeinsamer Zugriffspunkt*: Sowohl aus Anbieter- als auch aus Anwendersicht existiert ein einziges System zur Veröffentlichung und Suche von Web Services. Aus technischer Sicht werden dadurch Abhängigkeitsbeziehungen zwischen mehreren Nutzungs-Frameworks und mehreren Angebots-Frameworks für Web Services auf ein Minimum reduziert. Sowohl zur Suche als auch zur Veröffentlichung von Web Service-Angeboten wird dadurch ebenfalls die Anzahl der notwendigen Kommunikationsverbindungen minimiert.
- *Vollständigkeit von Angebotsinformationen*: Durch die Verwaltung aller Web Service-Angebote durch ein einziges System wird der Aufwand zur Erzeugung vollständiger Suchergebnisse verringert, da nur eine Datenbasis existiert, die bei einer Suche nach Angeboten zu berücksichtigen ist.

Trotz dieser Vorteile ergeben sich durch eine zentralisierte Verwaltung gleichzeitig folgende Nachteile:

- *Single-Point-of-Failure*: Der Verzeichnisdienst stellt in einer zentralisierten Infrastruktur eine Komponente dar, die bei einem Ausfall den Komplettausfall des Gesamtsystems bzw. des Marktes nach sich ziehen kann [PKY03]. Weiterhin ist auch die eingeschränkte Skalierbarkeit zentralisierter Systeme in diesem Zusammenhang als Nachteil zu nennen [Be02].

⁴ Diese Informationen dienen zur Einordnung von Anbieterunternehmen in eine Klassifikationsstruktur. Dazu gehört beispielsweise der Sitz oder der Geschäftsbereich eines Unternehmens.

⁵ Diese Informationen beziehen sich auf die Beschreibung von Web Services auf technischer Ebene. Dazu zählen beispielsweise, eine Schnittstellenspezifikation oder die URL, d.h. der Ort eines Web Services.

- *Angebotsmonopol*: Da es für die Dienstleistungen zur Verwaltung und Suche von Web Service-Angeboten nur einen einzigen Anbieter gibt, stellt der Betrieb eines Verzeichnisdienstes im Sinne einer Marktform ein Monopol dar. Dies führt dazu, dass die verwaltende Instanz des Verzeichnisdienstes bei der Gestaltung der Preise zur Dienstleistung nur auf die Nachfrage Rücksicht nehmen muss, ohne dabei die Angebote anderer Anbieter berücksichtigen zu müssen.

Bei einer Betrachtung einer solchen Verwaltungsinstanz als ein Unternehmen ist davon auszugehen, dass ein solcher Monopolist durch seine Bestrebung zur Gewinnmaximierung in der Lage ist, die Nachfrage zur Nutzung des Verzeichnisdienstes allein durch seine Preisgestaltung zu regulieren. Die Preise zur Suche bzw. zur Veröffentlichung von Web Service-Angeboten sind deshalb unabhängig von den Kosten, die zur Erbringung Dienstleistungen eines Verzeichnisdienst-Betreibers anfallen. Die zu erwartenden Wirtschaftlichkeitswerte für die Nutzung und das Angebot von Web Service-basierten Anwendungssystemen bzw. von Teilsystemen sind dadurch ebenfalls abhängig von der Preisgestaltung des Betreibers des Verzeichnisdienstes. Die Praxistauglichkeit des Gesamtkonzepts wird deshalb durch die Abhängigkeit der Verfügbarkeit mindestens eines Verzeichnisdienstes maßgeblich von Preisregulierungsmaßnahmen zur Nutzung dieser zentralen Dienste beeinflusst. Marktmodelle, bei denen eine Abhängigkeit zu einem Monopolisten besteht, erfordern deshalb den Eingriff einer unabhängigen Instanz, wie z.B. einem Staat, bei der Preisgestaltung.

Neben den Vor- und Nachteilen zentraler Verzeichnisdienste ist weiterhin auch das allgemeine Problem der potentiell mangelhaften Validität von Angebotsinformationen zu nennen. Durch die öffentliche Bereitstellung eines Verzeichnisdienstes und damit der Übertragung der Verantwortlichkeit zur Administration eigener Angebotsinformationen auf die Anbieter selbst, kann das Problem entstehen, dass die Einträge des Verzeichnisdienstes aus Aufwandsgründen nach der einmaligen Eintragung nicht aktualisiert werden. Der Erfolg von Bestrebungen zur öffentlichen Bereitstellung von UDDI-Registries ist u. a. deshalb in der Vergangenheit ausgeblieben. Grundsätzlich sind deshalb sowohl bei zentralisierten als auch bei dezentralisierten Konzepten zur Verwaltung von Angebotsinformationen mit Hilfe von Verzeichnisdiensten Mechanismen zur automatisierten Angebotsaktualisierung zu berücksichtigen. Beispielsweise kann durch das Monitoring von Web Services die Validität von Angebotsinformationen regelmäßig automatisch überprüft werden [Ru06]. Darauf aufbauend können veraltete Einträge aus dem Verzeichnis entfernt werden.

Ausgehend von den Problemstellungen, die aus dem konventionellen zentralisierten Charakter der Infrastruktur von Web Service-basierten Anwendungssystemen resultieren, werden in diesem Beitrag alternative Konzepte zur dezentralen Organisation von UDDI-Registries beschrieben und evaluiert. Abschnitt 3 widmet sich dabei den betrachteten Konzepten im Detail, ausgehend von der konventionellen Lösung zur Replikation von UDDI-Registries werden dazu ein Konzept zur dezentralen Organisation mit Hilfe eines Peer-to-Peer-Netzes und ein Konzept mittels UDDI-Crawlern vorgestellt. In Abschnitt 4 erfolgt eine Bewertung der Ansätze basierend auf einer Evaluation im Rahmen einer Referenzarchitektur eines Web Service-basierten Föderierten ERP-Systems

(ERP-Systems) [BM07]. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick auf weitere Forschungsarbeiten in Abschnitt 5.

3 **Vorstellung der betrachteten Konzepte**

Existierende Ansätze zur Verwaltung von UDDI-Registries haben aufgrund ihres zentralen Ansatzes sowohl Vor- als auch Nachteile, wobei letztere maßgeblich aus dem beschriebenen Single-Point-of-Failure und dem Angebotsmonopol resultieren.

In diesem Beitrag werden daher zwei alternative Ansätze mit alternativen Organisationskonzepten vorgestellt: erstens die dezentrale Organisation in einem Peer-to-Peer-Netz und zweitens mittels eines UDDI-Crawlers. Zuvor wird aber die konventionelle Lösung zur Replikation von UDDI-Registries noch einmal näher beleuchtet.

3.1 **Replikation von UDDI-Registries als konventionelle Lösung**

Zunächst ist an dieser Stelle festzuhalten, dass bereits Konzepte zur Verteilung von UDDI-Registries existieren, die sich auf die Replikation, d.h. den periodischen Abgleich, von UDDI-Datenbanken beziehen. In der UDDI-Spezifikation (Version 3) ist zu diesem Zweck die UDDI Replication API vorgesehen [Be04]. Die Verfügbarkeit der Funktionen eines Verzeichnisdienstes kann somit durch das redundante Angebot von Verzeichnisdienst-Instanzen verbessert werden. Eine verbesserte Skalierbarkeit von Publikations- und Suchdiensten ist ebenfalls gegeben. Der Vergleich der Kommunikationsbeziehungen zwischen den Teilnehmern eines Anwendungssystem-Netzwerks, unter Bezugnahme auf den Einsatz einer einzigen zentralen UDDI-Registry und der konventionellen Form zur Berücksichtigung mehrerer UDDI-Registry-Instanzen, ist in Abbildung 2 dargestellt.

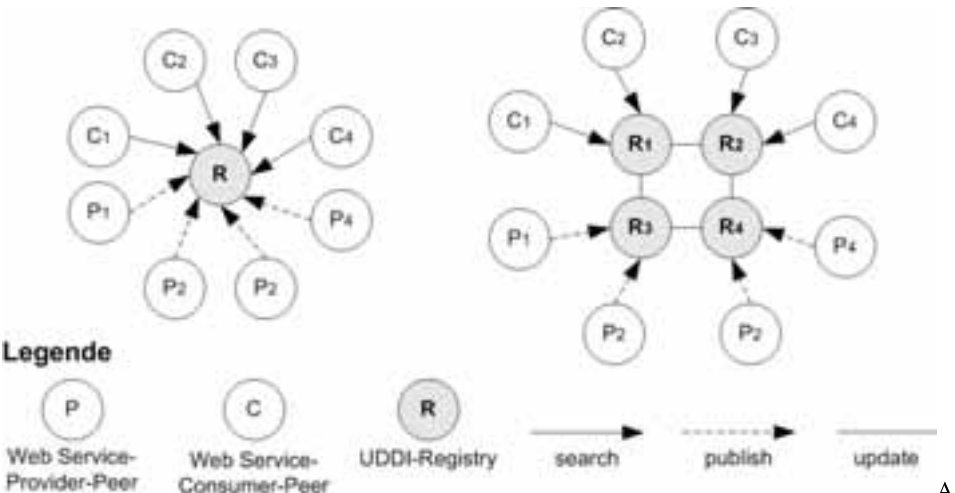


Abbildung 2: Vergleich zwischen singulärer UDDI-Registry und redundanter UDDI-Registry unter Nutzung von Replikationsmechanismen

Der durch die redundante Verfügbarkeit erreichbare Vorteil der Gleichheit des Datenbestandes aller UDDI-Datenbanken im Netzwerk, kann jedoch für den Anwendungsfall eines marktoffenen Web Service-basierten Anwendungssystems auch als Nachteil interpretiert werden⁶.

Zwar ist der Betrieb einzelner UDDI-Registry-Instanzen durch verschiedene Unternehmen denkbar, wodurch das Angebotsmonopol aufgehoben werden würde, jedoch verschwinden durch die zwanghafte Replizierung eigener Angebotsdatenbestände die Möglichkeiten zur Differenzierung des Angebots gegenüber konkurrierenden Verzeichnisdienst-Anbietern, da alle Anbieter nicht nur den gleichen Dienstyp, sondern auch annähernd den gleichen Dienst bereitstellen⁷. Die Bereitstellung privater UDDI-Registries⁸ kann diesem Problem zwar entgegenwirken, wodurch sich aber wiederum der Aufwand zur Einholung einer vollständigen Angebotsliste aus der Sicht eines Konsumenten erhöht, da dieser verschiedene UDDI-Registries kontaktieren muss. Weiterhin ist auch das Problem der automatisierten Benachrichtigung aller potentiellen Konsumenten eines neuen Such- und Publikationsdienstes zu nennen, was aus der Sicht des Betreibers einer UDDI-Registry einen zusätzlichen Aufwand verursacht.

Aus diesen Gründen sollen zwei alternative Konzepte vorgeschlagen werden, die auf die Bereitstellung von UDDI-Registries durch die Web Service-Anbieter selbst abzielen, wobei jeder Web Service-Anbieter die Menge seines eigenen Angebots in seiner eigenen UDDI-Registry verwaltet. Die Implementierungsergebnisse beider Ansätze werden nachfolgend in Kapitel 4 vorgestellt und evaluiert.

3.2 Dezentrale Organisation von UDDI-Registries in einem Peer-to-Peer-Netz

Ein alternativer Ansatz zum Verzicht auf eine zentrale UDDI-Registry, durch die Bereitstellung von UDDI-Registries durch die Web Service-Anbieter selbst, basiert auf der Integration von Web Service-Nutzern und -Anbietern durch ein Peer-to-Peer-Netzwerk. Das Peer-to-Peer-Netzwerk besteht dabei aus zwei unterschiedlichen Knoten⁹ (Peer)-Typen (Anbietern und Nutzern von Web Services), die jedoch gemäß der Definition eines Peer-to-Peer-Systems prinzipiell gleichberechtigt sind [Mi02]. Alle Peers verfolgen im gegebenen Kontext das gemeinsame Ziel der Bereitstellung und Aufrechterhaltung einer Infrastruktur zur Organisation von Web Service-Angeboten. In Abbil-

⁶ Weitere Nachteile im Bezug auf die mangelhafte Effizienz dieser Lösung sowie das entstehende Problem der Aktualität bzw. der „Richtigkeit“ von Einträgen in öffentlichen UDDI-Datenbanken wird in der Arbeit von [DHL06] diskutiert.

⁷ Unterschiede sind dabei lediglich im Bezug auf QoS-Eigenschaften, wie z.B. die Antwortzeiten unterschiedlicher Verzeichnisdienst-Instanzen möglich.

⁸ Hiermit sind UDDI-Registries gemeint, die nicht mit anderen UDDI-Registries verbunden sind.

⁹ Es wird darauf hingewiesen, dass verschiedene Ansätze zur Strukturierung von Peer-to-Peer-Netzwerken existieren, für die jedoch im gegebenen Kontext prinzipiell die gleichen Annahmen getroffen werden können.

dung 3 ist ein solcher Ansatz (rechts) im Vergleich zur Verwendung einer zentralen UDDI-Registry (links) dargestellt.

Die Darstellung zeigt deutlich die erhöhte Kompliziertheit der Netzwerkstruktur. Das Konzept basiert auf der Festlegung, dass alle Peers (Web Service Anbieter und Nutzer) über das Netzwerk miteinander verbunden sind. Jeder Anbieter verfügt über eine eigene UDDI-Registry, in der ausschließlich die eigenen angebotenen Web Services publiziert werden. UDDI-Suchanfragen werden unter Berücksichtigung des *Flooded Request Models* [Bu06] von einem Web Service-Konsumenten an alle bekannten Kontakte (Peers) gesendet¹⁰, die anschließend wiederum an die Kontakte der Kontakte weitergeleitet werden usw.

Ein Web Service-Anbieter, der eine Anfrage erhält, leitet die Suchanfrage an seine lokale UDDI-Registry weiter und sendet das Anfrageergebnis an den Initiator des Suchprozesses. Die Pfeile in Abbildung 3, die sich auf die Suche beziehen, kennzeichnen in diesem Zusammenhang die Kontakte eines Peers, an die eine Suchanfrage gesendet werden kann. Die dabei entstehende Struktur stellt ein Beispiel für diese Beziehungen dar.

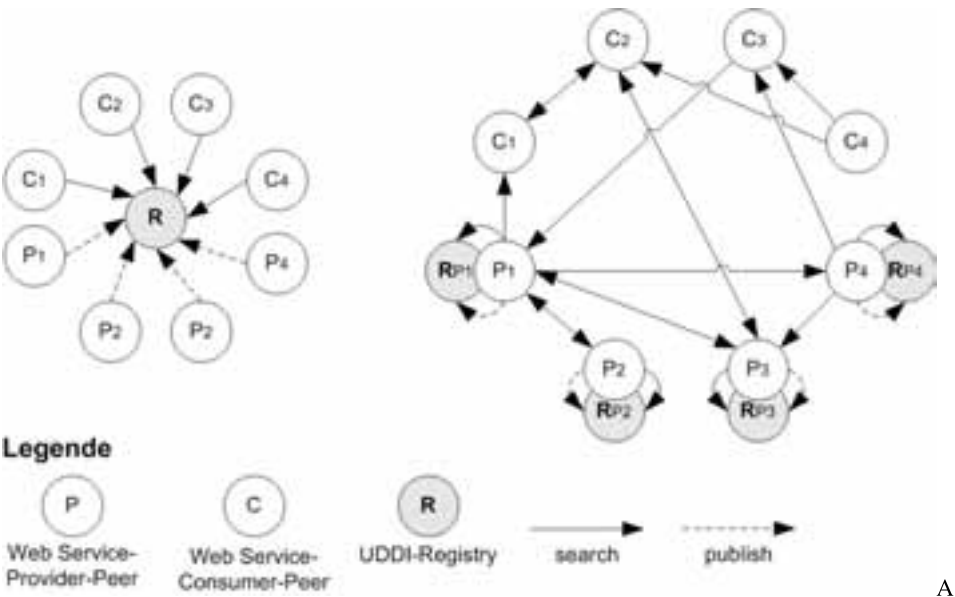


Abbildung 3: Vergleichende Darstellung zwischen einer zentralen und einer dezentralen FERP-Infrastruktur auf Basis eines Peer-to-Peer-Netzwerks

¹⁰ Hierbei kann es sich sowohl um andere Konsumenten als auch um Anbieter handeln.

Ein prinzipiell ähnlicher Ansatz wird von Papazoglou et al. beschrieben, wobei sich diese Ähnlichkeit auf die Weiterleitung von Anfragen bezieht, die nicht von einem Peer beantwortet werden können [PKY03]. Der Ansatz beschreibt weiterhin eine Möglichkeit zur Strukturierung von Angeboten bezogen auf Themengebiete, wie z.B. auf den Finanzbereich oder den Bereich des Marketings, die durch Zusammenschlüsse (Syndications) von Peers organisiert werden.

Der Ansatz basiert jedoch auf der Benachrichtigung von Peers im Fall der Änderung des Web Service-Angebots. Dazu wird ein Publish/Subscribe-Muster [PKY03] verwendet, bei dem alle interessierten Peers über neue Angebote bzw. über die Aktualisierung von Angeboten informiert werden. Bei einem Vergleich mit einem solchen Ansatz ist das in Abbildung 3 dargestellte Modell, als vereinfachte Variante zu betrachten.

Im Rahmen der Evaluation der Referenzarchitektur eines Web Service-basierten Föderierten ERP-Systems (FERP-Systems)¹¹ [BM07] wurde das in Abbildung 3 vorgeschlagene Modell prototypisch implementiert. Dazu wurde auf die verfügbaren Komponenten des JXTA-Projekts¹² zur Standardisierung von Peer-to-Peer-Anwendungen zurückgegriffen. Für die Implementierung solcher Anwendungen stehen Open-Source-Programm-bibliotheken u. a. auch für Java bereit.

Weiterhin existiert bereits ein Peer-to-Peer-Netzwerk, dem beigetreten werden kann, wodurch die Verfügbarkeit von Initialisierungsdiensten bereits gesichert ist. Zusätzlich werden durch das JXTA-Framework bereits alle notwendigen Funktionen zum Betrieb eines Peer-to-Peer-Netzwerks implementiert. Für eine detaillierte Darstellung der Funktionsweise und der Architektur wird beispielhaft auf die Arbeit von [QN04] verwiesen. Für einen Test wurde eine eigene Peer-Group mit dem Namen „FERPNet“ eingerichtet. Zur Umsetzung des Konzepts wurden folgende architekturbezogene Erweiterung der FERP-Referenzarchitektur durchgeführt:

- Das Basissystem zur Web Service-Nutzung wurde um die Teilkomponente *JXTAPeer* erweitert, die zur Teilnahme am JXTA-Netzwerk und im Speziellen an der Peer-Group „FERPNet“ dient. Diese Komponente exportiert eine Schnittstelle, die von einer Teilkomponente zur Erstellung von Suchanfragen importiert wird, so dass UDDI-Suchanfragen alternativ an diese Teilkomponente weitergeleitet werden können. Eine Initialisierung des Basissystems zur Web Service-Nutzung führt dadurch automatisch zur Anmeldung am JXTA-Netzwerk.

¹¹ Ein Föderiertes ERP-System (FERP-System) ist ein ERP-System dessen betriebliche Anwendungsfunktionen variabel, dynamisch oder adaptiv zu unterschiedlichen voneinander unabhängigen Softwareanbietern zugeordnet werden können. Die Gesamtfunktionalität des Systems wird von einem Ensemble standardisierter Teilsysteme bereitgestellt, das für den Endanwender als Gesamtsystem (ERP-System) erscheint [BM07]. Durch die in [BM07] dargestellte Referenzarchitektur wird die Bereitstellung von Fachkomponenten für ERP-Systeme auf Basis von Web Services vorgeschlagen. Die Referenzarchitektur beschreibt zunächst eine zentralisierte Infrastruktur, beispielsweise unter Nutzung einer zentralen UDDI-Registry.

¹² <https://jxta.dev.java.net/>

- Zur Vereinfachung der Gesamtproblematik wurde jedes Basissystem mit einer eigenen lokalen UDDI-Registry verbunden. Obwohl gemäß der Abbildung 3 Suchanfragen von Web Service-Anbietern beantwortet werden, die selbst jeweils eine lokale UDDI-Registry betreiben, bezieht sich die Evaluation des Testfalls auf die FERP Peer-to-Peer-Infrastruktur, wobei Anwenderunternehmen selbst Web Services für andere Anwenderunternehmen bereitstellen.

Ein Test der Funktionalität des Systems bestätigte zunächst Umsetzbarkeit des Konzepts einer FERP-Peer-to-Peer-Infrastruktur, die auf der FERP-Referenzarchitektur aufbaut. Bei einem Performanztest eines solchen JXTA-Netzwerks mit fünf Peers mit konstantem UDDI-Datenbestand wurde bei einer Variation der Anzahl sendender Peers sowie bei einer Variation der Anzahl versendeter Suchanfragen pro Peer durchgeführt. Als Ergebnis wurde eine nahezu konstante Antwortzeit zur Abfrage aller Peers ermittelt. Diese Antwortzeit betrug in etwa sechs Sekunden¹³ wobei die Anfragelast schrittweise auf insgesamt bis zu 120 Suchanfragen pro Minute pro Peer erhöht wurde.

Ein interessanter Aspekt ist weiterhin, dass eine maximale Anzahl von Anfragen zur fehlerfreien Anfragestellung ermittelt werden konnte, der bei ca. zehn Anfragen pro Sekunde liegt. Andere Versuche wie beispielsweise von Halepovic bestätigen dies, wobei ein Maximum von ca. 20 Nachrichten pro Sekunde ermittelt wurde [Ha04].

Die Ergebnisse zeigen, dass das vorgeschlagene Modell als Alternative zur Verwendung einer zentralen UDDI-Registry betrachtet werden kann. Als Nachteil dieses Modells ist zum einen die erhöhte Bearbeitungszeit einer Suchanfrage zu nennen, die auf Basis des beschriebenen Tests ermittelt wurde. Diesem Nachteil kann jedoch möglicherweise durch die lokale Archivierung von Suchergebnissen bei allen Web Service-Konsumenten in einem Puffer-Speicher (Cache) entgegengewirkt werden, so dass Daten über Web Service-Anbieter lediglich periodisch erneuert werden müssen. Als weiterer Nachteil ist der Aspekt zu nennen, dass aufgrund des dezentralen Charakters von Peer-to-Peer-Netzwerken, wobei ein Peer über keine direkte Verbindung mit allen anderen Peers verfügt, in der Regel eine Kooperationspflicht für alle Peers besteht. Bezogen auf ein Web Service-basiertes Anwendungssystem, wie z.B. einem FERP-System, bezieht sich diese Kooperationspflicht auf die Weiterleitung von Suchanfragen, sowohl durch Konsumenten als auch durch Anbieter. Zur Entgegenwirkung des dadurch potentiell entstehenden Problems der Unvollständigkeit von Suchanfragen bei unkooperativem Verhalten sei auf die Integration des FairNet-Protokolls verwiesen [Bu06], was sich auf die Erkennung und „Bestrafung“ von unkooperativem Verhalten in Peer-to-Peer-Datenstrukturen bezieht.

¹³ Für den Versuch wurden PCs (Windows XP) mit einem Intel Pentium 2,8 GHz Prozessor und 1 GB Hauptspeicher verwendet. Die PCs waren über einen Switch miteinander verbunden. Die Bandbreite für alle Verbindungen betrug 100 MBit/s.

Trotz der Berücksichtigung dieses Protokolls bleibt jedoch der Nachteil bestehen, dass insbesondere bei Web Service-Konsumenten möglicherweise ein mangelnder Anreiz zur Weiterleitung von Suchanfragen anderer Konsumenten besteht, wodurch die Akzeptanz des Gesamtkonzepts insgesamt negativ beeinflusst werden würde. Dieses Problem ist dadurch begründbar, dass die Konsumenten keinen einfach zu bestimmenden Profit aus ihrer Kooperation ziehen können. Da ein Suchdienst nur dann benötigt wird, wenn bisher keine Informationen über die Anbieter eines Web Service-Typs existieren, oder wenn eine Aktualisierung des „Marktüberblicks“ erzeugt werden soll. Diese Aktion ist jedoch abhängig von der individuellen Konfiguration des Basissystems zur Web Service-Nutzung bei einem Anwenderunternehmen.

Als Beispiel für den möglichen Verzicht zur eigenen Benutzung eines Suchdienstes soll der ausschließliche Erwerb von Pauschallizenzen, wie z.B. Bezahlung pro Monat, benannt sein, wobei eine Bindung eines Anwenderunternehmens an einen einzigen Anbieter besteht und keine Auswahl getroffen werden muss.

Im Gegensatz dazu ist eine verpflichtende Weiterleitung von Suchanfragen aus Sicht von Web Service-Anbietern eher umsetzbar, da diese einen direkten finanziellen Profit aus dem Angebot von Web Services erzielen können, der ohne die Existenz des Peer-to-Peer-Netzwerks eingeschränkt wird. Bezogen auf das oben genannte Beispiel des JXTA-Frameworks kann als möglicher erweiterter Ansatz die Verwendung separater Peer-Groups genannt werden, wodurch Konsumenten und Anbieter von Web Services voneinander getrennt werden. Jedoch sind dazu vertiefende Untersuchungen notwendig, wobei die Anforderungen an ein konkretes Web Service-basiertes Anwendungssystem und dessen Teilnehmer zu berücksichtigen sind.

3.3 UDDI-Crawler

Eine weitere Alternative zum Verzicht auf eine zentrale UDDI-Registry, basiert prinzipiell ebenfalls auf dem oben benannten Konzept zur eigenständigen Verwaltung und Bereitstellung lokaler UDDI-Registries durch die Anbieter von Web Services selbst. Bezugnehmend auf die genannten Nachteile von Peer-to-Peer-Infrastrukturen aus der Sicht von Web Service-Konsumenten soll hier eine Alternative vorgeschlagen werden, die sich auf die Berücksichtigung von sekundären Suchdiensten bezieht. Eine Umsetzungsmöglichkeit eines solchen Konzepts ist in Abbildung 4 (rechts) im Vergleich zum Ursprungskonzept dargestellt (links). Das Modell sieht zunächst verschiedene UDDI-Registries vor, die von Web Service-Anbietern (Provider) betrieben werden, wobei analog zum oben dargestellten Peer-to-Peer-Konzept ebenfalls nur die eigenen Web Service-Angebote publiziert werden (primäre UDDI-Registries). Im Kern dieses Modells werden Suchdienste (UDDI-Crawler) berücksichtigt, die in einer eigenen sekundären UDDI-Registry eine Kopie aller Angebote der bekannten primären UDDI-Registries speichern. Als UDDI-Crawler wird hier ein System bezeichnet, dass periodisch primäre UDDI-Registries automatisch durchsucht und den Datenbestand in einer eigenen sekundären UDDI-Registry speichert bzw. aktualisiert. Ein verwandtes Konzept zur Überwachung verschiedener UDDI-Registries wird in der Arbeit von Du et al. mit dem Namen

Ad-UDDI bezeichnet [DHL06], wobei im Speziellen auf die Verbesserung der Aktualität von UDDI-Datenbeständen durch ein Monitoringsystem abgezielt wird¹⁴.

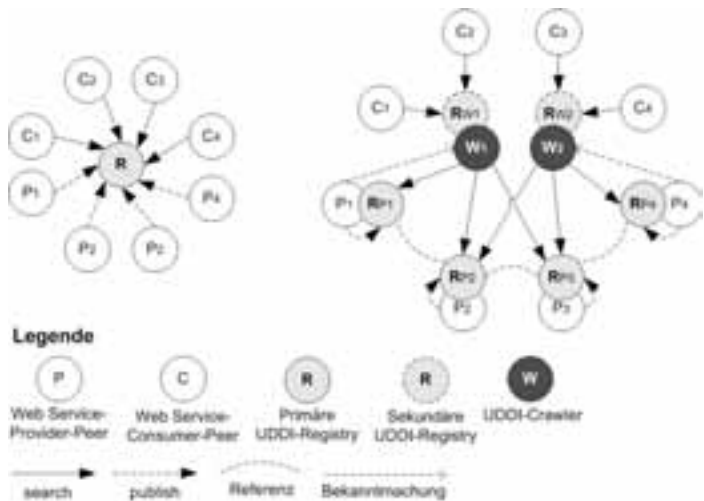


Abbildung 4: Vergleich zwischen einer zentralen UDDI-Registry und einem dezentralen System unter Nutzung von UDDI-Crawler-Instanzen

Das hier vorgestellte Konzept sieht zunächst eine Funktion zur Anmeldung primärer UDDI-Registries bei einem UDDI-Crawler vor. Diese Anmeldung wird von Web Service-Anbietern vorgenommen und ist mit der Anmeldung einer Unternehmenshomepage bei einer Suchmaschine für Webseiten vergleichbar. Da, wie in Abbildung 4 sichtbar, mehrere UDDI-Crawler existieren, die von unterschiedlichen Betreibern (Unternehmen) bereitgestellt werden können, ist auch eine Bekanntmachung bei mehreren UDDI-Crawler-Instanzen denkbar. Auch dies ist mit dem bereits bewährten Prinzip der Bekanntmachung einer Webseite bei verschiedenen Suchmaschinen vereinbar. Auf dieser Basis ist jedoch zunächst nur die Suche in primären UDDI-Registries möglich, die einem UDDI-Crawler bisher bekannt sind. Nicht berücksichtigt werden dabei Web Service-Anbieter, die ihre lokale UDDI-Registry nicht bei einem UDDI-Crawler angemeldet haben. Als weiteres Problem ist an dieser Stelle die Unvollständigkeit von Suchergebnissen zu nennen, die von sekundären UDDI-Registries an anfragende Web Service-Konsumenten zurückgeliefert werden, da nicht davon auszugehen ist, dass jeder Provider seine UDDI-Registry bei allen UDDI-Crawler-Instanzen bekanntgemacht hat. Dadurch ergibt sich das Problem des Ausschlusses existierender Angebote bei der Beantwortung von Suchanfragen. Auch dieses Problem kann mit Suchmaschinen für Webseiten verglichen werden, da nicht jeder Besitzer einer Homepage eine Anmeldung bei allen existie-

¹⁴ Ein anderes Konzept wird auch von Aiken dargestellt, wobei jedoch von einem UDDI-Crawler keine Kopien von Web Service-Angeboten verwaltet werden, sondern Anfragen transformiert und an UDDI-Registries weitergeleitet werden [Ai05]. In der Arbeit von Aiken, wird dieses Konzept ebenfalls mit dem Namen UDDI-Crawler bezeichnet.

renden Suchmaschinen vorgenommen hat und daher nicht in der Menge von Suchergebnissen berücksichtigt wird, die an einen Benutzer einer Suchmaschine gesendet werden.

Um diesem Problem entgegenzutreten, sieht das vorgeschlagene Konzept zusätzlich die Möglichkeit zur Referenzierung primärer UDDI-Registries untereinander vor, was zu einer Herausbildung eines UDDI-Registry-Netzes führen soll. Aufgabe eines UDDI-Crawlers ist es darauf aufbauend, eine erweiterte Suche durch das rekursive Durchsuchen und Kopieren des Datenbestands aller referenzierten UDDI-Registries durchzuführen. Auch an dieser Stelle ist ein Vergleich mit dem bewährten Konzept des World Wide Web (WWW) zulässig, da dieses Netz ebenfalls auf dem Prinzip der Referenzierung von Homepages untereinander basiert und von sog. Web-Crawlern durchsucht wird.

Die Motivation zur Referenzierung anderer Homepages ergibt sich dabei unter anderem aus den Nutzeffekten, die ein Unternehmen aus der Erweiterung des eigenen Informationsangebots erzielen kann. Beispielsweise kann der Betreiber einer Homepage zur Routenplanung auf die Homepage eines Hotels verweisen, dass sich nahe einer berechneten Route befindet. In diesem Zusammenhang sind Referenzen auf andere Homepages als „Bereicherung“ des eigenen Angebots zu betrachten. Bezogen auf den Anwendungsfall eines Web Service-basierten Anwendungssystems, wie einem ERP-System, kann sich ein ähnlicher Nutzen für einen Anbieter (A) aus der Referenzierung von UDDI-Registries anderer Web Service-Anbieter (B) ergeben, wenn im Angebot von Anbieter (B) ein Web Service-Typ enthalten ist, der die Voraussetzung für die Nutzung eines Web Services von Anbieter (A) ist. Bezogen auf ein ERP-System kann beispielsweise der Anbieter (A) eines Web Services zur Termin- und Kapazitätsplanung im Kontext der Produktionsplanung auf den Anbieter (B) eines Web Services zur Darstellung zeitbezogener Pläne in Form von Grafiken verweisen, da von einer kombinierten Verwendung beider Web Service-Typen aus der Sicht von Workflow-Designern auszugehen ist. Da die Auffindbarkeit eines konkreten Web Services des benannten Web Service-Typs zur Grafikerzeugung zur Darstellung von Planungsergebnissen aus Sicht von (A) essentiell ist, kann durch den Verweis auf das Angebot von (B) ein Nutzeffekt erzielt werden. Dieser Nutzeffekt ergibt sich aus einem potentiell erhöhten Konsum des Web Services von (A), da durch die Existenz des Web Services von (B) neben der „reinen“ Kapazitätsplanung und der direkten Übernahme von Plänen auch „Testplanungen“ im Sinne der Simulation fiktiver Kapazitäten und Zeiten möglich sind, die von der grafischen Darstellung von Planungsergebnissen für den Benutzer eines Web Service-basierten Anwendungssystems abhängig sind.

Zur Umsetzung des vorgeschlagenen Prinzips der Referenzierung von UDDI-Registries untereinander ist es möglich, UDDI-Registries selbst als Web Services in anderen Registries unter einer standardisierten Bezeichnung zu publizieren. Es gilt an dieser Stelle zu bemerken, dass UDDI ein SOAP-basierter Standard ist, was bedeutet, dass durch eine UDDI-Registry die Implementierung eines standardisierten Web Services dargestellt wird, wobei die gesamte Kommunikation auf der Basis des Austauschs von SOAP-Nachrichten erfolgt. Eine UDDI-Registry kann demnach auch als Web Service betrachtet werden, dessen Beschreibung (inkl. der URL) in anderen UDDI-Registries erfolgen kann. Dabei ist die URL zur sog. Inquiry-Schnittstelle von besonderem Interesse, da dies die Schnittstelle zur Abfrage des Datenbestands der referenzierten UDDI-Registry ist

und für einen UDDI-Crawler somit der Ausgangspunkt für einen Rekursionsschritt bei der Durchsuchung des gesamten Netzwerks darstellt.

Im Rahmen einer prototypischen Implementierung des vorgestellten UDDI-Crawler-Konzepts wurden folgende Festlegungen getroffen, die als Grundlage für eine Standardisierung des Referenzierungskonzepts im Kontext konkreter Anwendungssysteme dienen können:

- Es existiert in allen UDDI-Registries ein Eintrag des UDDI-Typs business-Entity mit dem Namen „UDDI-Network“.
- Alle Verweise auf andere UDDI-Registries werden diesem Eintrag untergeordnet und als Einträge vom UDDI-Typ businessService mit beliebigen Namen (z.B. Name des anderen Anbieters) eingefügt.
- Jeder businessService-Eintrag beinhaltet einen Eintrag vom UDDI-Typ description, der den Wert der Inquiry-URL der referenzierten UDDI-Registry enthält.

Unter den gegebenen Festlegungen wurde ein einfacher Test zur Bestimmung der Dauer der Kopie aller Einträge einer UDDI-Registry in eine andere durchgeführt, der sich auf die Abfrage aller Daten einer UDDI-Registry (A) und die anschließende Speicherung dieser Daten in UDDI-Registry (B) bezieht. Bei einer systematischen Erhöhung der Anzahl der Einträge in (A) wurden lineare Anstiege sowohl des zu übertragenden Datenvolumens als auch der Dauer des „Kopiervorgangs“ beobachtet. Die Größe der zu übertragenen Daten (HTTP-Pakete) betrug dabei durchschnittlich ca. 1 KByte pro Datensatz in (A). Unter Berücksichtigung ähnlicher Umgebungsbedingungen wie beim Test des Peer-to-Peer Netzwerks (siehe oben) wurde eine durchschnittliche Dauer des „Kopiervorgangs“ von etwa 50 ms pro Datensatz in (A) ermittelt. Als Wertebereiche für die bei den Tests berücksichtigten Datenbestände in (A) wurden Mengen zwischen 7.000 bis 140.000 Einträgen gewählt.

Unter der Annahme, dass die ermittelten Ergebnisse für alle Wertebereiche gelten, ist die periodische Aktualisierung des Datenbestands eines UDDI-Crawlers für Web Service-Anbieternetzwerke auch mit höherer Anzahl von Anbietern denkbar, da der Datenbestand einer primären UDDI-Registry möglicherweise nicht so hoch ist, wie dies im dargestellten Versuch angenommen wurde. Es gilt an dieser Stelle ebenfalls zu bemerken, dass der beschriebene Versuch auf die Datenübertragung in einem lokalen Netzwerk durchgeführt wurde, so dass keine Daten über das Internet versendet wurden, wie dies unter der in der Einleitung beschriebenen Motivation eines offenen Marktes für Web Service-basierte Anwendungssysteme der Fall wäre. Vertiefende Untersuchungen zur Bewertung der Praktikabilität des UDDI-Crawler-Konzepts sind deshalb unter Bezugnahme auf die tatsächliche Anzahl und den Typ angebotener Web Services notwendig und durch aufbauende Arbeiten zu leisten. Es bleibt jedoch festzuhalten, dass Umsetzung eines UDDI-Crawlers unter den benannten Festlegungen zur Referenzierung erfolgt ist, bei dem die Daten mehrerer UDDI-Registries als Kopie in einer sekundären UDDI-Registry gesammelt wurden. Auch die Ergebnisse des Performanztests für die Kopie einer UDDI-Registry deuten auf die mögliche Anwendung des Konzepts hin, wobei auch periodisch wiederholte „Kopiervorgänge“, z.B. einmal pro Tag, als durchführbar bewertet werden können.

4 **Gegenüberstellung der Ansätze**

Für einen Vergleich der Vor- und Nachteile der im den vorangegangenen Abschnitt vorgestellten Lösungsvorschläge zur Organisation von UDDI-Registries soll die Gegenüberstellung aus Tabelle 1 dienen. In der Tabelle ist zunächst der deutliche Nachteil des Ansatzes einer UDDI-Zentrale dargestellt, der sich insbesondere auf die mangelhafte Skalierbarkeit und die Fehlertoleranz im Fall eines Ausfalls der zentralen UDDI-Registry sowie die mangelhafte Marktoffenheit bezieht. Als besondere Vorteile sind die Performanz bei der Beantwortung von Suchanfragen, der Organisationsaufwand zur Aktualisierung von Web Service-Angeboten, d.h. insbesondere zur Bekanntmachung neuer Angebote, die Vollständigkeit von Suchergebnissen sowie die Aktualität des Angebotsdatenbestandes zu nennen¹⁵.

Ansatz	Skalierbarkeit Suchdienst	Performanz Antwortzeit der Suche	Organisationsaufwand Angebotsaktualisierung	Vollständigkeit Suchergebnisse	Aktualität Angebotsdatenbestand	Ausfalltoleranz Teil-Suchdienste	Marktoffenheit Suchdienste
Zentrale	- -	++	+++	+++	+++	- - -	- -
Replikation	++	++	-	+	+	+++	-
Peer-to-Peer	++	- -	+	+	+++	++	+
Crawler	++	++	- -	-	-	+	++

Tabelle 1: Gegenüberstellung der genannten Ansätze zur Organisation von UDDI-Registries

Bei der Betrachtung des Replikations-Konzepts sind als Vorteile die Performanz zur Beantwortung von Suchanfragen sowie die hohe Fehlertoleranz beim Ausfall einer UDDI-Registry-Instanz zu nennen. Nachteile des Ansatzes sind der Organisationsaufwand zur Aktualisierung aller UDDI-Registry-Instanzen sowie die eingeschränkte Marktoffenheit.

Das Peer-to-Peer-Konzept weist besondere Vorteile bezüglich der Skalierbarkeit von Suchdiensten, der Aktualität des Datenbestandes in UDDI-Registries sowie bei der Fehlertoleranz beim Ausfall einer UDDI-Registry-Instanz auf. Weiterhin ist die verbesserte Marktoffenheit des Konzepts zu nennen. Der grundlegende Nachteil wird insbesondere durch die reduzierte Performanz bei der Suche nach Web Service-Angeboten bestimmt.

¹⁵ Die Aktualität bezieht sich hierbei auf den Datenbestand, nicht auf die Korrektheit der angebotsbezogenen Daten, die bei allen Ansätzen von den Web Service-Anbietern selbst sicherzustellen ist.

Das UDDI-Crawler-Konzept ist bezüglich der Marktoffenheit zum Angebot von Suchdiensten als besonders vorteilhaft zu betrachten. Weiterhin sind die Eigenschaften der Skalierbarkeit und der Performanz bei der Suche von Vorteil. Als Nachteile sind der erhöhte Organisationsaufwand zur Aktualisierung des Datenbestands sekundärer UDDI-Registries, die mögliche Unvollständigkeit von Suchergebnissen sowie die Aktualität der Suchergebnisse zu nennen. Aufgrund des Vorteils der Marktoffenheit dieses Ansatzes wird eine weiterführende Betrachtung dieses Konzepts durch aufbauende Arbeiten mit höchster Priorität vorgeschlagen.

5 Zusammenfassung und Ausblick

Bei der Betrachtung von Web Services als vermarktbare Softwarekomponenten kommt der Verwendung von Verzeichnisdiensten essentielle Bedeutung zu, die allerdings nach wie vor mit Problemen belastet ist. Der Betrieb von UDDI-Registries als Beispiel für etablierte Verzeichnisdienste erfordert daher die ständige Weiterentwicklung und Bewertung der existierenden Konzepte mit dem Ziel, die Nachteile zu minimieren und sowohl Anbietern als auch Nutzern von Web Service-basierten Systemen Mehrwerte zu bieten. Dieser Beitrag stellte daher drei komplementäre Konzepte zum Betrieb von integrierten UDDI-Registries vor und bewertet sie bezüglich ihrer Vor- und Nachteile im Kontext eines offenen Web Service-Markts. Zur Evaluation wurden die Konzepte im Rahmen der Umsetzung eines Web Service-basierten Föderierten ERP-Systems (FERP-Systems) prototypisch implementiert.

In der Gegenüberstellung ergab sich dabei, dass grundsätzlich alle drei Ansätze, die konventionelle Lösung durch Replikation von UDDI-Registries ebenso wie die dezentrale Ansätze in einem Peer-to-Peer-Netz oder mittels eines UDDI-Crawlers gewisse Vor- und Nachteile aufweisen. Die Vorteile der dezentralen Ansätze liegen dabei aber insbesondere im Bereich der Marktoffenheit zum Angebot von Suchdiensten, also genau dort wo zentrale Ansätze bislang ein großes Manko aufweisen. Der vorgestellte Peer-to-Peer-Ansatz weist daneben trotz durchgehend positiver Aspekte (von Skalierbarkeit bis Aktualität und Ausfalltoleranz) insbesondere Mängel im Bereich der Performanz auf, die aus der erhöhten Bearbeitungszeit der Suchanfragen während der beschriebenen Tests resultiert. Der Crawler-basierte Ansatz weist diesen Nachteil nicht auf, erfordert aber einen höheren Organisationsaufwand zur Angebotaktualisierung durch den Einsatz von sekundären UDDI-Registries. Besonders im Hinblick auf die angestrebte Marktoffenheit ist dieser Ansatz aber eindeutig zu bevorzugen, da sekundäre UDDI-Registries in Zusammenhang mit UDDI-Crawlern von verschiedenen Anbietern unabhängig voneinander betrieben werden können.

Zukünftige Arbeiten könnten insbesondere im Bereich der Untersuchung zur Bewertung der Praktikabilität des UDDI-Crawler-Konzepts hinsichtlich Anzahl und Typ der von einem Provider angebotenen Web Services liegen. Darüber hinaus wäre eine Erweiterung des Peer-to-Peer-Ansatzes über das genannte Beispiel des JXTA-Frameworks in Bezug auf die Verwendung separater Peer-Groups möglich, um Anbieter und Konsumenten von Web Services voneinander zu trennen. Dabei sind die besonderen Anforderungen konkreter Anwendungssysteme zu berücksichtigen.

6 Literaturverzeichnis

- [Ai05] Aiken, D.: D10.1v0.1 Focused Crawler for Web Service Discovery. In (Kilgariff, E. Hrsg.): WSMX Working Draft, 2005.
- [Be02] Benatallah, B.; Dumas, M.; Sheng, Q.Z.; Ngu, A.H.H.: Declarative Composition and Peer-to-Peer Provisioning of Dynamic Web Services. 18th International IEEE Conference on Data Engineering ICDE. IEEE Computer Society, San Jose, USA, 2002; S. 297-308.
- [Be04] Bellwood, T.; Capell, S.; Clement, L.; Colgrave, J.; Dovey, M.J.; Feygin, D.; Hately, A.; Kochman, R.; Macias, P.; Novotny, M.; Paolucci, M.; von Riegen, C.; Rogers, T.; Sycara, K.; Wenzel, P.; Wu, Z.: UDDI Spec Technical Committee Draft. In (Clement, L.; Hately, A.; von Riegen, C.; Rogers, T. Hrsg.): OASIS, 2004, S. 18 ff.
- [BM07] Brehm, N.; Marx Gómez, J.: Web Service-basierte Referenzarchitektur für Föderierte ERP-Systeme. In (Pietsch, T.; Lang, C.V. Hrsg.): Ressourcenmanagement. Erich Schmidt Verlag, Berlin, 2007, S. 125-142.
- [BMZ08] Brehm, N.; Marx Gómez, J.; Ziesenitz, A.: Toolunterstützung bei der vermarktungsorientierten Entwicklung von Web Services als Bausteine komplexer betrieblicher Anwendungssysteme. In (Bichler, M.; Hess, T.; Krcmar, H.; Lechner, U.; Matthes, F.; Picot, A.; Speitkamp, B.; Wolf, P. Hrsg.): Multikonferenz Wirtschaftsinformatik 2008, München/Garching, GITO-Verlag, Berlin, 2008; S. 1853-1864.
- [BS03] Boles, D.; Schmees, M.: Kostenpflichtige Web Services. In (Uhr, W.; Esswein, W.; Schoop, E. Hrsg.): Wirtschaftsinformatik 2003: Märkte – Medien – Mobilität. Physica-Verlag, Dresden, 2003; S. 385-403.
- [Bu06] Buchmann, E.: Erkennung und Vermeidung von unkooperativem Verhalten in Peer-to-Peer-Datenstrukturen. Fakultät für Informatik, Vol. Dissertation. Otto-von-Guericke-Universität, Magdeburg, 2006.
- [DHL06] Du, Z.; Huai, J.; Liu, Y.: Ad-UDDI: An Active and Distributed Service Registry-Technologies for E-Services. Vol. 3811/2006. Springer, Berlin, Heidelberg, 2006; S. 58-71.
- [Ha04] Halepovic, E.: Performance Evaluation and Benchmarking of the JXTA Peer-To-Peer Platform. College of Graduate Studies and Research, Vol. Masters Thesis, University of Saskatchewan, Saskatoon, Kanada, 2004; S. 51.
- [Ha06] Hahn, H.: Die Etablierung von Märkten für Fachkomponenten. Eine vergleichende Analyse zur Beschaffung betrieblicher Anwendungssoftware und alternativer Güterkategorien. Wirtschaftsinformatik & Systems Engineering, Vol. Dissertation, Universität Augsburg, Augsburg, 2006.
- [HW02] Höß, O.; Weisbecker, A.: Konzeption eines Repositories zur Unterstützung der Wiederverwendung von Software-Komponenten. In (Turowski, K. Hrsg.): 4. Workshop Komponentenorientierte betriebliche Anwendungssysteme (WKBA 4), Augsburg, 2002; S. 67-85.
- [JT03] Jaekel, H.; Teschke, T.: Eine Infrastruktur für den Austausch von Fachkomponenten. In (Turowski, K. Hrsg.): 5. Workshop Komponentenorientierte betriebliche Anwendungssysteme (WKBA 5), Augsburg, 2003; S. 167-189.
- [Me99] Merz, M.: Elektronische Dienstmärkte – Modelle und Mechanismen des Electronic Commerce. Springer Verlag, Berlin, 1999.

- [Mi02] Milojevic, D.S.; Kalogeraki, V.; Lokuse, R.; Nagaraya, K.; Pruyne, J.; Richard, B.; Rollins, S.; Xu, Z.: Peer-to-Peer Computing. HP Labs, 2002; S. 4.
- [OT05] Overage, S.; Thomas, P.: WS-Specification: Ein Spezifikationsrahmen zur Beschreibung von Web-Services auf Basis des UDDI-Standards. In (Ferstl, O.K.; Sinz, E.J.; Eckert, C.; Isselhorst, T. Hrsg.): Wirtschaftsinformatik 2005: eEconomy – eGovernment – eSociety. Physica-Verlag, Bamberg, 2005; S. 1539-1558.
- [Ov04] Overage, S.: Die Spezifikation – kritischer Erfolgsfaktor der Komponentenorientierung. In (Turowski, K. Hrsg.): 4. Workshop Komponentenorientierte betriebliche Anwendungssysteme. Gesellschaft für Informatik: Arbeitskreis 5.10.3, Augsburg, 2004; S. 7 – 23.
- [PKY03] Papazoglou, M.P.; Krämer, B.J.; Yang, J.: Leveraging Web-Services and Peer-to-Peer Networks. In (Goos, G.; Hartmanis, J.; van Leeuwen, J. Hrsg.): Proc. 15th International Conference on Advanced Information Systems Engineering (CAiSE 2003), Vol. 2681, Klagenfurt/Velden, Österreich, Springer, Berlin, 2003.
- [QN04] Qu, C.; Nejdl, W.: Interacting the Edutella/JXTA Peer-to-Peer Network with Web Services. International Symposium on Applications and the Internet (SAINT'04), IEEE Computer Society, Tokyo, Japan, 2004; S. 67 – 73.
- [Ru06] Rud, D.: Qualität von Web Services. Messung und Sicherung der Performance. VDM Verlag Dr. Müller, Saarbrücken (2006)
- [Te03] Teschke, T.: Semantische Komponentensuche auf Basis von Geschäftsprozessmodellen. Department für Informatik, Vol. Dissertation. Universität Oldenburg, Oldenburg, 2003; S. 4.
- [ZK95] Zimmermann, H.-D.; Kuhn, C.: Grundlegende Konzepte einer Electronic Mall. In (Schmid, B.; Dravta, R.; Kuhn, C.; Mausberg, P.; Meli, H.; Zimmermann, H.-D. Hrsg.): Electronic Mall: Banking und Shopping in globalen Netzen, B. G. Teubner Verlag, Stuttgart, 1995.

Zentrale Gegenstände der Parametrisierung bei betrieblichen Softwarekomponenten

Jörg Ackermann, Klaus Turowski

Lehrstuhl für Wirtschaftsinformatik und Systems Engineering
Universität Augsburg
Universitätsstraße 16, 86135 Augsburg
{joerg.ackermann | klaus.turowski}@wiwi.uni-augsburg.de

Zusammenfassung. Mit dem Einsatz der Komponentenorientierung wird die Hoffnung verbunden, flexiblere und einfacher anzupassende betriebliche Anwendungssysteme erstellen zu können. Die Praxis zeigt, dass die dabei verwendeten Softwarekomponenten selbst auch anpassbar sein müssen. Parametrisierung ist ein bekanntes Verfahren, welches vor allem zur Anpassung der fachlichen Funktionalität eingesetzt wird. Diese Arbeit beschäftigt sich mit den Grundlagen parametrisierbarer betrieblicher Softwarekomponenten. Dazu werden die zentralen Gegenstände der Parametrisierung (Parameter, Parametergruppen, Parametrisierungsaufgaben und -auswirkungen) bestimmt, präzise definiert und deren Eigenschaften diskutiert. Darüber hinaus wird ein Parameter-Metamodell vorgestellt.

1 Einleitung

Erfolg und Misserfolg von Unternehmen werden zunehmend von deren Fähigkeit bestimmt, schnell auf veränderte Marktsituationen reagieren zu können. Bei den von Unternehmen derzeit eingesetzten betrieblichen Informationssystemen handelt es sich jedoch häufig um große, integrierte Anwendungen, die nur schwierig zu warten [Tu03:1] und nur schwer an veränderte Anforderungen anzupassen sind [Sz02:5].

Komponenten- und serviceorientierte betriebliche Anwendungssysteme haben demgegenüber den Vorteil, dass ein Teil der notwendigen Flexibilität durch Auswahl und Austausch von Komponenten bzw. Services erreicht werden kann. Allerdings kann ein Austausch allein nicht die gesamte notwendige Variabilität zur Verfügung stellen und ist bei häufigen, kleineren Änderungen auch nicht effizient, weil dieser mit einem nicht unerheblichen Aufwand verbunden sein kann. Eine Alternative dazu besteht im Einsatz parametrisierbarer Softwarekomponenten. So lässt sich eine notwendige Variabilität bei komponentenorientierten betrieblichen Anwendungen dann gut durch Parametrisierung realisieren, wenn eine lokal begrenzte Funktionalität anzupassen ist, sich häufige Änderungen ergeben oder mehrere Varianten gleichzeitig zum Einsatz kommen sollen [AT06:357]. Daraus ergibt sich die Motivation zur näheren Untersuchung parametrisierbarer betrieblicher Softwarekomponenten.

Zur Parametrisierung von Komponenten existieren verschiedene Ansätze, die sich vor allem in der Art der Ablage der Parameterwerte unterscheiden: Eingesetzte Ablagemechanismen sind Datenbanktabellen, einfache Initialisierungsdateien sowie die von den gängigen Komponententechnologien vorgesehenen XML-Konfigurationsdateien (Property-Files beim OMG CORBA Component Model [OMG02:1-47], Deployment-Descriptor bei Suns Enterprise Java Beans (EJB) [Mo00:234], Datei Web.config bei Microsofts .NET [Pr02:377]). Die eingesetzten Parametrisierungskonzepte unterscheiden sich jedoch zum Teil deutlich in Bezug auf Funktionsumfang und verwendeter Terminologie – eine einheitliche, technologieübergreifende Betrachtung von Parametrisierung fehlt bisher.

Ziel dieser Arbeit sind daher die Bestimmung und die umfassende Darstellung der für Parametrisierung relevanten Eigenschaften betrieblicher Softwarekomponenten. Dazu wird zunächst Parametrisierung als ein mögliches Anpassungsverfahren in komponentenorientierten Anwendungssystemen vorgestellt (Kapitel 2) und eine Beispielkomponente zur weiteren Illustration eingeführt (Kapitel 3). Danach wird ein Schema hergeleitet, das die Bestimmung der zentralen Gegenstände der Parametrisierung erlaubt (Kapitel 4). Die so identifizierten Gegenstände werden in den nachfolgenden Kapiteln ausführlicher diskutiert: es werden Parameter und Parametergruppen (in Kapitel 5), Parametrisierungsaufgaben (in Kapitel 6) und Parametrisierungsauswirkungen (in Kapitel 7) definiert und deren zentrale Eigenschaften anhand von Klassifikationsschemata herausgearbeitet. Außerdem wird ein Parameter-Metamodell vorgestellt. Die Arbeit schließt mit dem zusammenfassenden Kapitel 8.

Die originären Beiträge dieser Arbeit liegen in der systematischen Bestimmung der zentralen Gegenstände der Parametrisierung sowie in der Definition und der ausführlichen Charakterisierung dieser Gegenstände. Nach Wissen der Autoren wurden die zentralen Parametrisierungseigenschaften bei betrieblichen Fachkomponenten noch nie und allgemein bei betrieblichen Anwendungssystemen noch nicht so umfassend und systematisch herausgearbeitet. Der hier entwickelte Begriffsrahmen kann damit helfen, weitergehende Untersuchungen zur Parametrisierung zu strukturieren und wurde in diesem Sinne schon bei der Erarbeitung einer Spezifikationsmethodik für parametrisierbare Fachkomponenten eingesetzt [AT07].

2 Parametrisierung zur Anpassung von Softwarekomponenten

Anpassung spielt eine wichtige Rolle in komponentenorientierten Anwendungssystemen, da sich in der Praxis gezeigt hat, dass Komponenten nur selten ohne Anpassung wiederverwendet werden können [Bo97:13]. Dabei kann es sich entweder um eine Anpassung der Komponentenarchitektur sowie Auswahl und Zusammenspiel der einzelnen Komponenten oder um eine Anpassung der Komponenten selbst handeln [SC98:305]. Gängige Anpassungsmechanismen sind: Kopieren von Code, Vererbung, Aggregation, Verwendung von Wrapperkomponenten, Superimposition, Adapterinterfaces, parametrisierte Verträge und Parametrisierung – für einen Vergleich siehe z. B. [Bo97] oder [Re01:48-55]. Für einen allgemeinen Überblick zur Anpassung in komponentenorientierten Anwendungssystemen siehe z. B. [Re01].

Bei der Anpassung einzelner Komponenten kann man zwischen geplanter und ungeplanter Anpassung unterscheiden [Ac04:132]. *Geplante Anpassung* bedeutet dabei, dass die Anpassungsmöglichkeiten vom Komponentenhersteller vorgesehen werden. Bei betrieblichen Fachkomponenten wird darüber hinaus zwischen technischer und fachlicher Anpassung unterschieden [Tu03:44]. Die *technische Anpassung* dient dazu, implementierungsbedingte, technische Inkompatibilitäten zu beheben. Unter *fachlicher Anpassung* werden Tätigkeiten verstanden, die die betriebswirtschaftlichen, aufgabenbezogenen Eigenschaften einer Komponente ändern.

Diese Arbeit beschäftigt sich mit Anpassung durch Parametrisierung. Unter *Parametrisierung* (im engeren Sinne) wird ein Verfahren zur Anpassung von Softwarekomponenten (oder anderen Softwareeinheiten) verstanden, bei dem der Komponentenhersteller Parameter vordefiniert und der Komponentenanwender die Parameter mit geeigneten Werten belegt (vgl. auch Kapitel 5). Für die Parameter wird dabei von einer Datenfeldpragmatik ausgegangen, d. h. die Parameter werden mit nicht-ausführbaren Daten belegt. Unter nicht-ausführbaren Daten werden solche Daten verstanden, die kein Programm darstellen (wie z. B. Skripte) und von keiner anderen Seite als Programm interpretiert werden (wie z. B. Workflowschemata). Man kann in diesem Zusammenhang auch von *datenbasierter Parametrisierung* sprechen [Ac04:134]. Zur Abspeicherung der gewählten Parameterwerte können Datenbanktabellen oder die von den gängigen Komponententechnologien vorgesehenen Konfigurationsdateien (z. B. Property Files beim OMG CORBA Component Model [OMG02:1–47] oder Deployment Descriptors bei Suns Enterprise Java Beans (EJB) [Mo00:234]) verwendet werden. Es bleibt zu erwähnen, dass datenbasierte Parametrisierung – außerhalb komponentenorientierter Systeme – vielfach in kommerziell vertriebener Software (wie z. B. in betriebswirtschaftlicher Standardsoftware oder Office-Produkten) zum Einsatz kommt.

Der Begriff *Parametrisierung* wird auch außerhalb des Bereichs Softwarekomponenten verwendet, bezeichnet dabei aber zum Teil Konzepte, die sich bei aller Ähnlichkeit in einigen Punkten unterscheiden: In der Wirtschaftsinformatik wird *Parametrisierung* häufig synonym zu *Customizing* verwendet, worunter die Anpassung einer Standardsoftware an die Anforderungen eines Kunden verstanden wird [Gö97:101]. Ein prominentes Beispiel dafür ist das komplexe, parametergetriebene Customizing von SAP R/3 [SAP02]. Im Software Engineering ist *Parametrisierung* als eine Implementierungstechnik bekannt, die im Zusammenhang mit Wiederverwendung und Variabilität eingesetzt wird. Anwendungsgebiete finden sich im Reuse-driven Software Engineering Business (RSEB) [JGJ97], bei der generischen Programmierung [CE00] und bei Produktlinienansätzen [SB00].

3 Beispielkomponente *Lagermanagement*

Zur Illustration der weiteren Ausführungen wird in diesem Abschnitt eine Beispielkomponente *Lagermanagement* eingeführt. Um ein möglichst realistisches Beispiel zu erhalten, orientieren sich Design und Terminologie der Komponente an der betriebswirtschaftlichen Standardsoftware SAP R/3 Enterprise [SAP02] sowie an einem anerkannten Referenzdatenmodell [BS04]. Um die Verständlichkeit des Beispiels zu unterstützen,

wurde die Komponente jedoch deutlich vereinfacht – reale Anwendungen sind typischerweise komplexer. Die betriebliche Aufgabe der Komponente besteht darin, einen einfachen Lagerkomplex zu verwalten. Die von der Komponente verwalteten Daten werden in Abb. 1 durch ein Datenmodell dargestellt. Die Komponente erlaubt die Definition mehrerer *Lager*. Jedes Lager besteht aus einer Anzahl von *Lagerplätzen*, auf denen die Materialien physisch gelagert werden. Eine Instanz von *Lagerbestand* repräsentiert eine Einheit eines Materials, welche auf einem bestimmten Lagerplatz gelagert wird. Der Typ *Material* enthält die lagerspezifischen Eigenschaften eines Materials. Zur Vereinfachung wird angenommen, dass jedes Material in genau einem Lager gelagert wird – auf die Verwendung komplexer Lagerfindungsstrategien wird verzichtet. Außerdem werden im Beispiel keine Lagereinheitentypen verwendet und es wird stattdessen angenommen, dass innerhalb eines Lagers jeder Lagerplatz für ein Material geeignet ist.

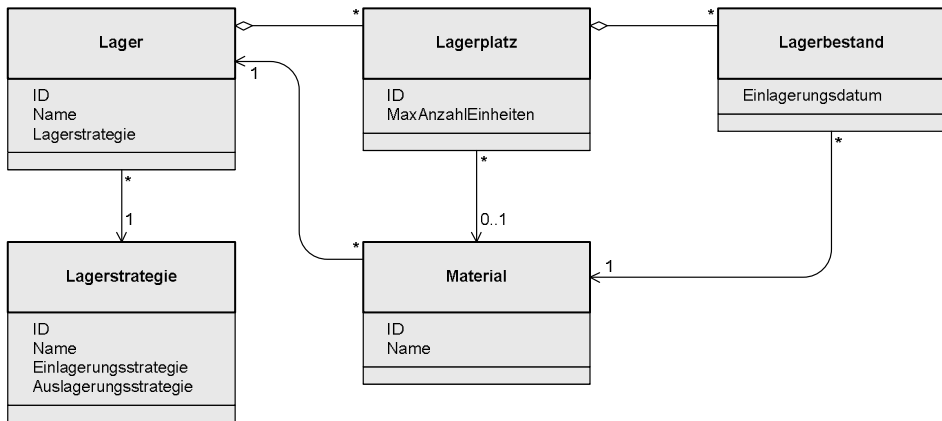


Abbildung 1: Von der Komponente *Lagermanagement* verwaltete Daten

Außerdem erlaubt die Komponente mehrere Varianten der Lagersteuerung einzusetzen (z. B. Fixplatzlager oder Hochregallager). Dazu enthält der Typ *Lagerstrategie* neben einer ID und einem Namen die zwei Steuerungsparameter *Einlagerungsstrategie* und *Auslagerungsstrategie*, mit denen die anzuwendenden Verfahren für Einlagerung (z. B. *dynamisch* oder *statisch*) und Auslagerung (z. B. *First-In-First-Out*) auszuwählen sind. Jedem Lager ist die für das Lager geeignete Lagerstrategie zuzuordnen.

Die Komponente bietet Schnittstellen zur Parametrisierung (z. B. *ILagerstrategie*) und zum Aufruf der betrieblichen Dienste der Komponente (z. B. *ILagermanagement*) an – auf detailliertere Ausführungen dazu wird an dieser Stelle verzichtet.

4 Ermittlung der zentralen Gegenstände der Parametrisierung

Zum besseren Verständnis der parametrisierungsrelevanten Eigenschaften betrieblicher Softwarekomponenten werden in diesem Beitrag die zentralen Gegenstände der Parametrisierung vorgestellt und beschrieben. Unter dem Begriff *zentrale Gegenstände der Parametrisierung* werden dabei alle wesentlichen Konzepte und Begriffe verstanden, die

im Zusammenhang mit Parametrisierung stehen. In diesem Kapitel werden dazu zunächst die zentralen Gegenstände bestimmt und überblicksartig vorgestellt.

Bei der Untersuchung von Parametrisierung ist es hilfreich, nach Sachverhalten zu unterscheiden, die zur Anpassungszeit und zur Laufzeit von Bedeutung sind [Ac04:136]:

- Der Begriff *Parametrisierungsoptionen* bezeichnet die zur Anpassungszeit relevanten Eigenschaften zur Parametrisierung und fasst dabei die funktionalen und nicht-funktionalen Eigenschaften der vom Hersteller vorgesehenen Parameter sowie die Möglichkeiten zur Parameterpflege zusammen.
- Unter *Parametrisierungsauswirkung* wird eine Variabilität an Datenstruktur oder Verhalten der Software verstanden, welche durch das Setzen von Parameterwerten erreicht werden kann und damit zur Laufzeit relevant ist.

Beispiele für Parametrisierungsoptionen und -auswirkungen finden sich in der Beispielkomponente *Lagermanagement* (vgl. Kapitel 3): *Einlagerungsstrategie* ist ein obligatorischer Parameter, der durch die Schnittstelle *ILagerstrategie* gesetzt wird und dessen Pflege aus fachlicher Sicht durch die Aufgabe *Einlagerungsstrategie festlegen* beschrieben wird. Eine Auswirkung des Parameters *Einlagerungsstrategie* besteht darin, dass er festlegt, mit welcher Strategie die betriebliche Aufgabe der Einlagerung ausgeführt wird.

Anzumerken bleibt, dass die Trennung in Parametrisierungsoptionen und -auswirkungen in der Praxis so nicht möglich ist, da für einen Komponentenanwender die Parametrisierungsoptionen immer nur im Zusammenhang mit den daraus resultierenden Auswirkungen interessant sind. Diese Trennung ist jedoch für die weiteren Untersuchungen sinnvoll, da sie den Gesamtkomplex der Eigenschaften zur Parametrisierung unterteilt und dadurch ein strukturiertes Vorgehen ermöglicht.

Zur Bestimmung der zentralen Gegenstände der Parametrisierung spielen zwei Aspekte eine wesentliche Rolle:

- Es ist (wie zuvor diskutiert) nach Parametrisierungsoptionen und Parametrisierungsauswirkungen zu unterscheiden.
- Zur Strukturierung von Anwendungssystem(teil)en wird häufig zwischen den Sichten für Daten, Funktionen und Prozessen unterschieden: Eine solche Strukturierung (mit teils abweichenden Begriffen) findet sich z. B. bei der Modellierung betrieblicher Anwendungssysteme [Sc98:37; SAP97], bei der Entwicklung komponentenorientierter (und anderer) Anwendungssysteme [DW98:43], zur Strukturierung der Außensicht von Softwarekomponenten [Ov06:126] sowie als Modellperspektiven bei der Systemtheorie [Ro99:75-81].

	Datensicht	Funktionssicht	Prozesssicht
Parametrisierungs- optionen	Parameter und Parametergruppen	Parametrisierungs- aufgaben	Reihenfolge von Parametrisierungs- aufgaben
Parametrisierungs- auswirkungen	Auswirkungen auf betriebliche Begrif- fe und Daten	Auswirkungen auf betriebliche Aufga- ben und Funktionen	Auswirkungen auf Reihenfolge von Aufgaben und Funktionen

Abbildung 2: Bestimmung der zentralen Gegenstände der Parametrisierung

Durch diese zwei Aspekte ergibt sich die Struktur der Matrix in Abb. 2, welche die zentralen Gegenstände der Parametrisierung enthält. Die so identifizierten Gegenstände werden ebenfalls in Abb. 2 überblicksartig aufgeführt, allerdings erst in späteren Kapiteln hergeleitet und detaillierter beschrieben: Für die Parametrisierungsoptionen werden zunächst in Kapitel 5 Parameter und Parametergruppen (Datensicht) und dann in Kapitel 6 Parametrisierungsaufgaben und deren Reihenfolge (Funktions- und Prozesssicht) untersucht. Danach werden die laufzeitrelevanten Parametrisierungsauswirkungen im Kapitel 7 thematisiert.

5 **Parameter und Parametergruppen**

Die zentralen Gegenstände bei den Parametrisierungsoptionen aus Datensicht sind naturgemäß gerade die Parameter und deren Struktur. In den folgenden Unterabschnitten werden die Parameter selbst sowie ihre strukturelle Zusammenfassung zu Parametergruppen näher untersucht. Die so gewonnenen Erkenntnisse werden danach in einem Parameter-Metamodell strukturiert abgebildet.

5.1 **Parameter**

Es gibt derzeit keine einheitliche Definition des Begriffs *Parameter* für die Anpassung von Softwarekomponenten. Ausgehend von bestehender Literatur (vgl. [AT06:343f.] für eine ausführliche Diskussion) werden die Begriffe *Parameter* und *Parameterwert* wie folgt definiert:

„Ein (*datenbasierter*) *Parameter* ist ein Datenfeld, welches zu einer Softwarekomponente (oder zu einem anderen Softwarebaustein) gehört und sich durch folgende Merkmale auszeichnet:

- Das Datenfeld dient zur geplanten Anpassung der Software und wird dazu vom Hersteller der Software zur Entwicklungszeit vordefiniert. Der Hersteller definiert neben dem Datenfeld auch dessen Bedeutung und dessen Auswirkungen auf Struktur und Verhalten der Software.
- Der Verwender der Software belegt das Datenfeld zur Anpassungszeit so mit Werten, dass die Software seinen Anforderungen entsprechend arbeitet.
- Die vom Verwender ausgewählten Werte sind nicht-ausführbar, werden (z. B. in Datenbanktabellen oder Dateien) abgespeichert und zur Laufzeit ausgewertet.“

„*Parameterwert* bezeichnet einen Wert, der einem Parameter aktuell zugewiesen ist.“

Es sei darauf hingewiesen, dass ein Parameter nach dieser Definition ein nicht weiter strukturiertes Datenfeld ist. Komplexere Parameterstrukturen sind natürlich bei der Parametrisierung möglich, werden aber als Parametergruppe und nicht als Parameter bezeichnet. Der Grund dafür liegt – neben der Anlehnung an die Literatur – darin, dass damit eine begriffliche Trennung erfolgt zwischen (den für die Steuerung einer Komponente verantwortlichen) elementaren Parametern und der Struktur der Parameter.

Es bleibt anzumerken, dass sich Parameter zur fachlichen Anpassung betrieblicher Anwendungssysteme teilweise nur schwer von Anwendungsdaten (insbesondere Stammdaten) abgrenzen lassen [Gö97:102, AR00:65]. Ansatzpunkte für eine Unterscheidung ergeben sich dadurch, dass Parameter im Gegensatz zu Anwendungsdaten

- einem Anwender einen Entscheidungsspielraum im Sinne alternativer Eingabewerte bieten (im Gegensatz zu nicht beeinflussbaren Ist-Werten) [DMH99:3],
- typischerweise von fachlichen Systemverantwortlichen (Superuser) und nicht vom Endanwender gesetzt werden [Gr98:481],
- zur Konfigurationszeit zu pflegen sind (spätere Änderungen sind meist möglich) [Ri00:258],
- bei Softwarekomponenten oft über eigene Schnittstellen zu ändern sind, die nur für Konfigurationszwecke und nicht durch andere Softwarekomponenten des Anwendungssystems aufrufbar sein sollten [OMG02:1-45].

Im Sinne dieser Unterscheidung handelt es sich (bei der Beispielkomponente aus Kapitel 3) bei Lagersteuerung um Parameterdaten und bei Lagerplätzen und Materialien um Stammdaten. Zu beachten ist, dass damit (in Übereinstimmung mit anderen Autoren [Gö97:101, SGR98:152, DFH00:535]) organisatorische Einheiten (wie z. B. Lager) ebenfalls zu den Parameterdaten gezählt werden.

Zum besseren Verständnis werden wichtige Eigenschaften von Parametern (bei betrieblichen Softwarekomponenten) mit Hilfe des Klassifikationsschemas in Abb. 3 dargestellt. Die im Schema enthaltenen Merkmale stammen aus der Literatur zur Parametrisierung [Pi93; Gö97; DMH99] sowie aus einer Analyse des Customizings von SAP R/3 [Ac02].

Merkmal	Merkmalsausprägung			
Datentyp	Boolesch	Numerisch	Alpha-numerisch	Zeichenartig
Wertebereich	Beliebig	Eingeschränkt	Festwerte	Parameterabhängig
Notwendigkeit	Optional		Bedingt optional	Obligatorisch
Defaultwert	Ja		Nein	
Charakter	Identifizierend	Beschreibend	Referenzierend	Steuernd

Abbildung 3: Klassifikationsschema für Parameter

Das Klassifikationsschema in Abb. 3 enthält folgende Merkmale:

- Das Merkmal *Datentyp* erfasst qualitativ, welche Art von Datentyp der Parameter hat: Nach [DMH99:4] oder [Ac04:137] sind Parameter immer formatiert und treten in den Zeichenarten *boolesch*, *numerisch*, *alphanumerisch* und *zeichenartig* auf.
- Das Merkmal *Wertebereich* beschreibt die Freiheiten bei der Wahl von Parameterwerten [Ac02:28; Pi93:439]. Zu unterscheiden ist zwischen folgenden Ausprägungen: im Rahmen seines Datentyps nicht eingeschränkt (*Beliebig*), mit Einschränkungen versehener Standarddatentyp (*Eingeschränkt*), vom Hersteller fest vorgegeben (*Festwerte*) oder durch die Werte anderer Parameter vorgegeben (*Parameterabhängig*).
- Das Merkmal *Notwendigkeit* gibt an, ob ein Parameter mit einem Wert versehen werden muss [Ac02:28]. Mögliche Ausprägungen sind *Obligatorisch*, *Optional* und *Bedingt optional*. Letzteres bedeutet, dass der Parameter nur unter bestimmten Bedingungen (meist andere Parametereinstellungen) optional ist.
- Das Merkmal *Defaultwert* gibt an, ob für den Parameter eine Voreinstellung vorgesehen ist, die bei fehlender Wertzuweisung verwendet wird – ein initialer Wert ohne semantische Bedeutung wird dabei nicht als Defaultwert betrachtet.
- Das Merkmal *Charakter* unterscheidet Parameter nach ihrer Rolle bei der Anpassung (einzelne Merkmalsausprägungen gehen auf [Ac02:28f.], [Ac04:147] und [Gö97:102] zurück): Von zentraler Bedeutung sind *steuernde* Parameter, welche unmittelbar Datenstrukturen, Funktionsabarbeitung oder Prozessabläufe beeinflussen. So definiert z. B. der Parameter *Einlagerungsstrategie* die bei der

Einlagerung anzuwendende Strategie. Mit Hilfe von *identifizierenden* Parametern kann zwischen verschiedenen Parametrisierungsvarianten unterschieden werden – ein Beispiel dafür ist der Parameter *Id* der Parametergruppe *Lagersteuerung*. *Beschreibende* Parameter legen z. B. Mengeneinheiten oder Texte fest, haben aber keine steuernde Funktion – so erlaubt der Parameter *Name* der Parametergruppe *Lager* eine nähere Beschreibung eines Lagers. Ein Parameter einer Parametergruppe hat *referenzierenden* Charakter, wenn er auf eine Instanz einer anderen Parametergruppe verweist und damit Eigenschaften der ersten Gruppe festlegt. Der Parameter *Strategie* von *Lager* weist einem Lager eine Lagerstrategie zu, welche die konkreten Steuerungseigenschaften für dieses Lager festlegt.

Einlagerungsstrategie ist ein Beispiel für einen Parameter bei der Komponente *Lagermanagement*. Der Datentyp des Parameters ist zeichenartig und sein Wertebereich ist durch Festwerte vorgegeben (*manuell*, *statisch*, *dynamisch*). Die Pflege des Parameters ist obligatorisch, es gibt keinen Defaultwert und sein Charakter ist steuernd.

5.2 Parametergruppen

Die Parameter einer Komponente treten häufig nicht isoliert auf, sondern liegen in strukturierter Form vor – für solche Struktureinheiten wird ein eigener Begriff eingeführt: Eine *Parametergruppe* ist eine Struktur semantisch zusammengehörender Parameter, zu welcher mehrere Instanzen (Datensätze) angelegt werden können. Ein wesentlicher Aspekt dieser Definition ist die Forderung, dass eine Parametergruppe mehrere Instanzen enthalten und damit als Träger verschiedener Anpassungsvarianten dienen kann.

Ein Beispiel für eine Parametergruppe ist *Lagersteuerung*, welche z. B. die Parameter *Einlagerungsstrategie* und *Auslagerungsstrategie* enthält. Es ist möglich, verschiedene Instanzen von *Lagersteuerung* zu definieren, um unterschiedliche Lager in verschiedener Weise steuern zu können.

Welche Parameter zu einer Parametergruppe zusammengefasst werden, ist nicht immer eindeutig zu beantworten und liegt daher teilweise im Ermessensspielraum des Komponentenherstellers. Ein Kriterium dafür liegt vor, wenn die Instanzen einer Parametergruppe durch einen semantischen Schlüssel eindeutig identifiziert werden: Zur Parametergruppe können nur solche Parameter gehören, welche dem Entitätstyp zugeordnet sind, der durch den Schlüssel beschrieben wird. Werden durch den Schlüssel z. B. Lager identifiziert, dann können zur Parametergruppe nur Parameter gehören, die eine Steuerung auf Lagerebene erlauben und nicht solche, die eine Steuerung auf Lagerbereichsebene intendieren.

Merkmal	Merkmalsausprägung			
Bwl. Zweck	Organisatorische Einheiten	Strategische Steuerdaten	Administrative Steuerdaten	Benutzersteuerung/Darstellung
Max. Anzahl Instanzen	Eine	Feste Anzahl	Parameterabhängig	Beliebig

Abbildung 4: Klassifikationsschema für Parametergruppen

Wichtige Eigenschaften von Parametergruppen finden sich in Abb. 4:

- Das Merkmal *Betriebswirtschaftlicher Zweck* beschreibt die Funktion der Parametergruppe auf betriebswirtschaftlicher Ebene und ihren Einfluss auf die Systemsteuerung. Dieses Merkmal wird nicht in eine formale Spezifikation eingehen, da insbesondere die Abgrenzung zwischen den Merkmalsausprägungen nicht eindeutig ist. Als Zusatzinformation für den menschlichen Leser ist dieses Merkmal jedoch eine sinnvolle Ergänzung. Bei diesem Merkmal werden folgende (nicht notwendig disjunkte) Merkmalsausprägungen unterschieden [Gö97:101f.; Ac02:26]:
 - Bei der Parametrisierung werden *organisatorische Einheiten* (z. B. Werke, Lager) definiert, welche einerseits Datenhaltungsebenen sind [DFH00:535] und andererseits häufig mittelbaren Einfluss auf die Ablaufsteuerung haben, indem sie die Träger verschiedener Ablaufvarianten sind.
 - Steuerdaten steuern den Ablauf von Prozessen und Geschäftsvorfällen. Dabei definieren *strategische Steuerdaten* abstrakte Regeln und Strukturen für Prozessabläufe. *Administrative Steuerdaten* steuern konkrete Prozessabläufe und sind oft Ausprägungen der strategischen Steuerdaten.
 - Unter *Benutzersteuerung/Darstellung* werden Parametergruppen zusammengefasst, welche Darstellung und Bildschirmdetails sowie Benutzerbesonderheiten steuern, aber keinen Einfluss auf den Ablauf von Prozessen haben.

- Das Merkmal *Maximale Anzahl Instanzen* beschreibt, wie viele Datensätze für die Parametergruppe maximal definiert werden können [Ac02:27]. *Feste Anzahl* heißt, die maximale Anzahl ist vom Komponentenhersteller (z. B. durch Festwerte) vorgegeben. *Parameterabhängig* heißt, die maximal erlaubte Anzahl wird durch Parameterwerte oder durch die Anzahl von Instanzen anderer Parametergruppen festgelegt.

Lagersteuerung ist ein Beispiel für eine Parametergruppe bei der Beispielkomponente *Lagermanagement* – diese enthält z. B. den Parameter *Einlagerungsstrategie*. Diese Parametergruppe gehört zu den administrativen Steuerdaten, welche die Lagerprozesse steuern. Die Komponente ermöglicht, beliebig viele Instanzen von *Lagersteuerung* zu definieren.

5.3 Parameter-Metamodell

In diesem Abschnitt wird ein Parameter-Metamodell vorgestellt, welches eine Weiterentwicklung des Modells aus [Ac04:139] ist. Ein solches Metamodell erlaubt, die zuvor definierten Begriffe rund um Parameter weiter zu präzisieren, indem sie strukturiert und mit ihren Abhängigkeiten untereinander dargestellt werden. Bei der Erstellung des Metamodells wurde die Struktur von Parametern in Datenbanktabellen, in einfachen Initialisierungsdateien sowie in XML-Konfigurationsdateien (Property-Files beim OMG CCM [OMG02:1–47], Deployment-Descriptor bei Suns EJB [Mo00:234], Datei Web.config bei Microsofts .NET [Pr02:377]) berücksichtigt. Im Ergebnis entstand ein Metamodell, welches für alle Spielarten der datenbasierten Parametrisierung gültig ist. Zu beachten ist dabei, dass das Metamodell in dem Sinne ein Maximalmodell ist, dass nicht alle dargestellten Konzepte in allen Spielarten der Parametrisierung vorkommen müssen.

Das Metamodell wird als UML-Klassendiagramm [OMG05:21] dargestellt und findet sich in Abb. 5. Im Metamodell werden Typ- und Instanzebene unterschieden. Die Elemente der Typebene werden vom Komponentenhersteller vorgegeben und umfassen die verfügbaren Parameter sowie deren Strukturierung in Parametergruppen:

- Zentrales Element des Metamodells ist der Parameter, worunter ein Datenfeld verstanden wird, welches bei der Parametrisierung zum Einsatz kommt. Ein Parameter wird durch seinen Namen sowie seine Eigenschaften Datentyp, Wertebereich, Notwendigkeit und Defaultwert näher beschrieben (vgl. dazu Abb. 3). Bei datenbankbasierter Ablage entspricht ein Parameter einer Spalte einer Tabelle – bei XML-Konfigurationsdateien handelt es sich bei den Parametern um die Elemente (Knoten unterster Ebene) oder die Attribute.
- Parametergruppen werden im Metamodell durch den Typ *Parametergruppe* repräsentiert. Eine Parametergruppe wird durch ihren Namen näher beschrieben und kann beliebig viele Parameter umfassen. Bei datenbankbasierter Ablage entspricht eine Parametergruppe einer Tabelle. Bei XML-Konfigurationsdateien bilden alle die Knoten eine Parametergruppe, die andere Knoten oder Elemente enthalten. Das Metamodell erlaubt auch Parameter ohne Parametergruppe. Während dies bei datenbankbasierter Ablage unüblich ist, tritt dieser Fall bei

Dateien dann auf, wenn ein Element auf höchster Ebene definiert wird. Dieser Fall kommt z. B. bei einfachen Initialisierungsdateien häufig vor.

- Parametergruppen können hierarchisch angeordnet sein. Dies wird im Metamodell durch die reflexive Beziehung bei *Parametergruppe* beschrieben. Hierarchien von Parametergruppen entstehen insbesondere bei XML-Konfigurationsdateien dann, wenn mehrstufige Knotenhierarchien verwendet werden.

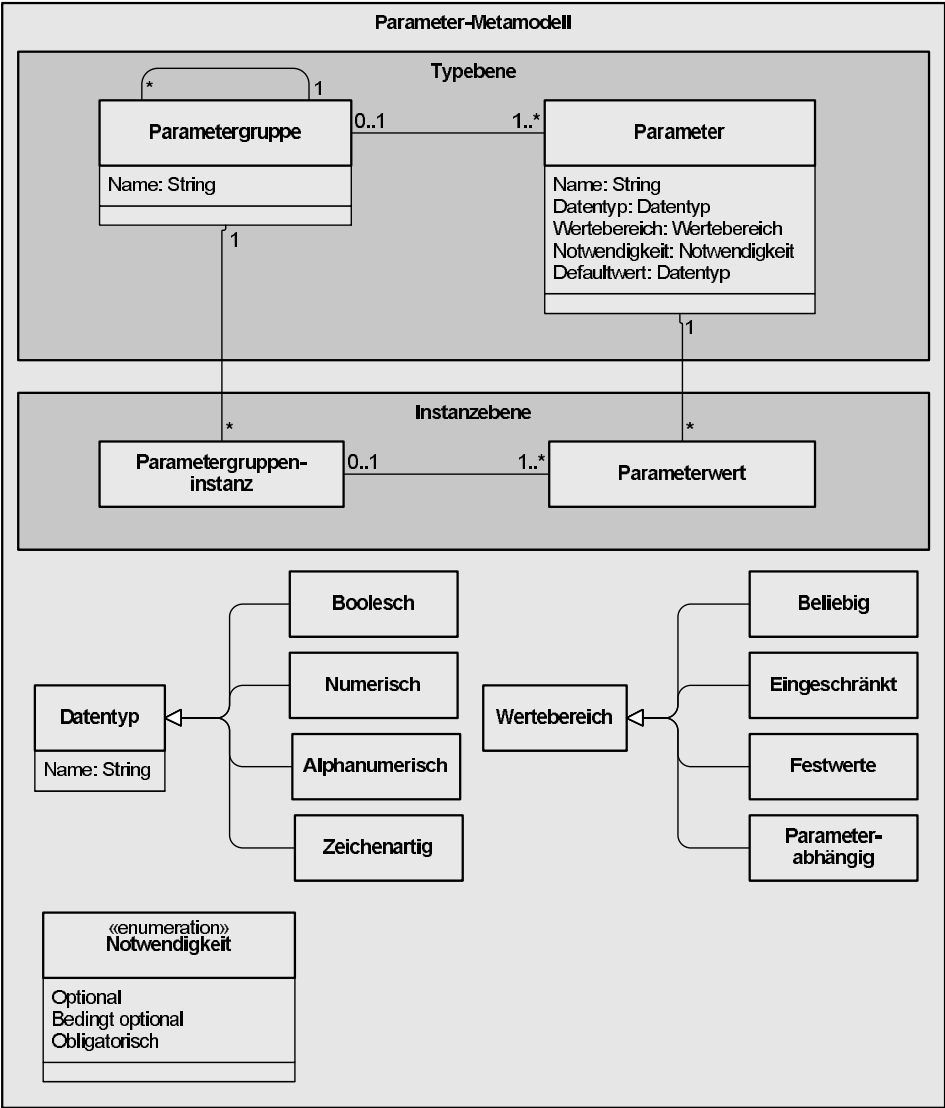


Abbildung 5: Metamodell zur Beschreibung von Parametern

Auf der Instanzebene finden sich die Daten, welche ein Komponentenanwender den Parametern zuweist. In Entsprechung zur Typebene werden dabei unterschieden:

- *Parameterwert* ist ein einem Parameter zugewiesener Wert, d. h. der Inhalt eines Datenbankfeldes bzw. der Inhalt eines Elements oder Attributs in einer XML-Konfigurationsdatei. Ein Parameterwert bezieht sich immer auf einen Parameter und gehört zu einer oder keiner Parametergruppeninstanz. Parameterwerte müssen den Restriktionen entsprechen, die durch Datentyp und Wertebereich ihres Parameters vorgegeben werden.
- Eine Parametergruppe kann mit beliebig vielen Ausprägungen versehen werden, was durch die *Parametergruppeninstanz* repräsentiert wird. Eine Parametergruppeninstanz (auf unterster Hierarchiestufe) ist dabei ein Tuple von Werten, welches aus genau einem Parameterwert für jeden Parameter der Parametergruppe besteht. Bei datenbankbasierter Ablage entspricht dies gerade einem Datensatz in der Tabelle. In XML-Konfigurationsdateien handelt es sich um das Vorkommen eines Knotens inklusive seiner enthaltenen Attribute, Knoten, und Elemente. Das Anlegen mehrerer Instanzen zu einer Parametergruppe ermöglicht die Definition verschiedener Parametrisierungsvarianten, die z. B. für unterschiedliche organisatorische Einheiten gelten können.

Ein Parameter bei der Beispielkomponente *Lagermanagement* ist die für ein Lager zu definierende *Einlagerungsstrategie*, welche den Festwert *dynamisch* als möglichen Parameterwert hat. Dieser Parameter ist Teil der Parametergruppe *Lagersteuerung*. Eine Parametergruppeninstanz ist z. B. eine Lagersteuerungsinstanz mit der Id 02, für welche als Einlagerungsstrategie *dynamisch* und als Auslagerungsstrategie *First-In-First-Out (FIFO)* festgelegt wird.

6 Parametrisierungsaufgaben

6.1 Definition und Beschreibung

In Abb. 2 sind der Funktionssicht bei den Parametrisierungsoptionen solche Gegenstände zuzuordnen, mit deren Hilfe das Erfassen von Parameterwerten durch einen Komponentenanwender ermöglicht wird. Auf technischer Ebene sind das z. B. spezielle Schnittstellen zur Pflege von Parameterwerten. Auf konzeptioneller Ebene gibt es dafür bisher keinen eindeutigen Begriff – in der Literatur wird meist nur der gesamte Prozess der Anpassung bezeichnet (z. B. als Customizing [Gö97:101; DMH99:1; SGR98:152], Parametr(is)ierung [AR00:64; Me05:165], Adaption [Hu94]) oder allgemein von Parametereinstellungen [Me05:167] gesprochen.

Um die Funktionen zur Parametererfassung auf konzeptioneller Ebene zu repräsentieren, wird – analog zum Begriff der betrieblichen Aufgabe – der Begriff der *Parametrisierungsaufgabe* eingeführt:

- Eine *elementare Parametrisierungsaufgabe* repräsentiert eine Handlung, bei der eine Menge semantisch zusammengehörender Parameter mit einem eindeutig abgegrenzten, nicht mehr unterteilbaren Parametrisierungsziel, eingestellt werden.
- Eine *zusammengesetzte Parametrisierungsaufgabe* beschreibt ein übergeordnetes Parametrisierungsziel und ist dazu aus einer Menge anderer Parametrisierungsaufgaben zusammengesetzt.

Ein wesentlicher Aspekt bei diesen Definitionen ist die Unterscheidung in elementare und zusammengesetzte Parametrisierungsaufgaben: Um ein Parametrisierungsziel einfach beschreiben zu können und Abhängigkeiten zwischen Parametrisierungszielen klar erfassen zu können, müssen Parametrisierungsaufgaben so feingranular wie möglich sein – dafür gibt es die elementaren Parametrisierungsaufgaben. Zusammengesetzte Parametrisierungsaufgaben dagegen ermöglichen eine Strukturierung von elementaren und anderen zusammengesetzten Parametrisierungsaufgaben, die einem übergeordneten Parametrisierungsziel dienen.

Für Beispiele wird wieder auf die Komponente *Lagermanagement* zurückgegriffen: Die bei der betrieblichen Aufgabe *Einlagerung* einzusetzende Strategie wird durch den Parameter *Einlagerungsstrategie* der Parametergruppe *Lagersteuerung* festgelegt – die Pflege dieses Parameters erfolgt mit Hilfe der elementaren Parametrisierungsaufgabe *Einlagerungsstrategie festlegen*. Analog dazu bestimmt die Parametrisierungsaufgabe *Auslagerungsstrategie festlegen* die bei der Auslagerung einzusetzende Strategie. Eine Zusammenfassung der Parameter *Ein-* und *Auslagerungsstrategie* zu einer elementaren Parametrisierungsaufgabe ist nicht sinnvoll, da beide Entscheidungen unabhängig voneinander sind und verschiedene betriebliche Aufgaben betreffen. Die o. g. Parametrisierungsaufgaben lassen sich mit der Aufgabe *Lagersteuerung definieren* zur zusammengesetzten Parametrisierungsaufgabe *Lagersteuerung einrichten* zusammenfassen.

Durch das Konzept der elementaren Parametrisierungsaufgabe erfolgt eine andere Sichtweise auf die in der Datensicht eingeführten Parameter und Parametergruppen. Eine Parametrisierungsaufgabe kann zur Pflege eines einzelnen oder mehrerer Parameter dienen und erzeugt dadurch eine Gruppierung von Parametern. Diese zusammen zu pflegenden Parameter können zu verschiedenen Parametergruppen gehören und müssen auch nicht alle Parameter einer Parametergruppe umfassen. Es ist außerdem denkbar, dass ein Parameter von verschiedenen Parametrisierungsaufgaben eingestellt wird. Es besteht also keine eindeutige Zuordnung einer Parametrisierungsaufgabe zu einem einzelnen Parameter oder zu einer Parametergruppe. Dies ist dadurch begründet, dass die Zusammenfassung von Parametern zu Parametergruppen strukturorientiert und zu Parametrisierungsaufgaben zielorientiert (nach Wirkungsweise) erfolgt.

Merkmal	Merkmalsausprägung			
Parametrisierungskontext	Aufbauorganisation	Daten	Funktionen	Abläufe (Prozesse)
Notwendigkeit	Optional	Bedingt optional		Obligatorisch
Abhängigkeiten	Keine	Komponentenlokal		Komponentenübergreifend

Abbildung 6: Klassifikationsschema für Parametrisierungsaufgaben

Wichtige Eigenschaften von Parametrisierungsaufgaben werden wieder in einem Klassifikationsschema aufgeführt (vgl. Abb. 6). Das Schema enthält folgende Merkmale:

- Das Merkmal *Parametrisierungskontext* beschreibt, welche Aspekte eines Anwendungssystems durch eine Parametrisierungsaufgabe angepasst werden (Merkmalsausprägungen *Aufbauorganisation*, *Daten*, *Funktionen* und *Abläufe*) [AT06:352].
- Das Merkmal *Notwendigkeit* gibt an, ob eine Parametrisierungsaufgabe ausgeführt werden muss oder nicht. Obligatorische Aufgaben müssen in jedem Fall vor Einsatz der Komponente durchgeführt werden – optionale dagegen nicht unbedingt. *Bedingt optional* bedeutet, dass die Aufgabe unter bestimmten Voraussetzungen (häufig aufgrund anderer Parametereinstellungen) obligatorisch wird.
- Das Merkmal *Abhängigkeiten* erfasst, ob zur Durchführung der Aufgabe die vorhergehende Ausführung anderer Parametrisierungsaufgaben notwendig ist [Ac02:26]. Bei bestehenden Abhängigkeiten wird dabei unterschieden, ob alle vorausgesetzten Parametrisierungsaufgaben zur selben Komponente gehören (*komponentenlokal*) oder teilweise anderen Komponenten zugeordnet sind (*komponentenübergreifend*).

Zur Erläuterung der Klassifikation wird die Parametrisierungsaufgabe *Einlagerungsstrategie festlegen* betrachtet: Anhand dieser Aufgabe wird festgelegt, nach welchem Verfahren die betriebliche Aufgabe der Einlagerung ausgeführt wird – Funktionen sind daher der Parametrisierungskontext. Die Aufgabe ist obligatorisch, da Einlagerung immer eingesetzt wird und damit immer einzustellen ist. Um die Aufgabe zu bearbeiten, muss zunächst die Parametrisierungsaufgabe *Lagersteuerung definieren* ausgeführt werden – es bestehen also komponentenlokale Abhängigkeiten.

6.2 Reihenfolgebeziehungen

In der Prozesssicht von Parametrisierungsoptionen werden Abläufe und Reihenfolgebeziehungen von Funktionen (bzw. Aufgaben) beschrieben. Die dafür relevanten Gegenstände der Parametrisierung zur Anpassungszeit sind einzuhaltende Abläufe und bestehende Reihenfolgerestriktionen bei der Abarbeitung von Parametrisierungsaufgaben. Ein Beispiel dafür findet sich bei der Beispielkomponente: Die Parametrisierungsaufgabe *Lagersteuerung definieren* ist Voraussetzung für die Aufgaben *Einlagerungsstrategie festlegen* und *Auslagerungsstrategie festlegen*.

Im Vergleich zu Geschäftsprozessen weisen die Abläufe zur Parametrisierung jedoch einen geringeren Umfang und eine geringere Komplexität auf. Dies erkennt man z. B. daran, dass bei SAP R/3 zwar die Geschäftsprozesse mittels Ereignisgesteuerter Prozessketten (EPK) modelliert werden, die Abläufe zur Parametrisierung jedoch nicht. Eine explizite Modellierung der Reihenfolgebeziehungen zwischen Parametrisierungsaufgaben in Form von Prozessen erscheint – unter Berücksichtigung des Grundsatzes der Wirtschaftlichkeit der Grundsätze ordnungsmäßiger Modellierung [BRS95:434] – nicht sinnvoll. Aus diesem Grund wird an dieser Stelle auch auf die explizite Einführung und Erläuterung eines eigenen Begriffs (wie z. B. *Parametrisierungsprozess*) verzichtet. Aber auch ohne eine Modellierung in Form von Prozessen bleibt zu betonen, dass Reihenfolgebeziehungen zwischen Parametrisierungsaufgaben bestehen und dass diese bei der Spezifikation einer betrieblichen Softwarekomponente zu berücksichtigen sind.

7 Parametrisierungsauswirkungen

Das Ziel von Parametrisierung ist, Struktur und Verhalten einer Softwarekomponente an die individuellen Anforderungen des Komponentenanwenders anzupassen. Die zur Anpassungszeit gesetzten Parameterwerte beeinflussen dabei zur Laufzeit die Datenstruktur sowie die Eigenschaften der ausgeführten Funktionen und Prozesse. Alle solche Änderungen werden unter dem Begriff *Parametrisierungsauswirkungen* zusammengefasst (vgl. Kapitel 4). Beispielsweise entscheidet der Parameter *Einlagerungsstrategie* (der Komponente *Lagermanagement*), mit welcher Strategie die betriebliche Aufgabe der Einlagerung ausgeführt wird (Auswirkung auf Funktionssicht).

Wirkungsart	Beschreibung der Wirkungsart
Varianten- bestimmend	Identifiziert eine anzuwendende Parametrisierungsvariante
Definierend	Definiert alle erlaubten Ausprägungen
Auswählend	Wählt aus vorgegebenen Ausprägungen die erlaubten aus
Einschränkend	Schränkt erlaubte Ausprägungen ein
Entscheidend	Trifft Entscheidungen bei Prozessablauf oder Funktionsabarbeitung
Entscheidungs- unterstützend	Unterstützt Entscheidungsfindung bei Prozessablauf oder Funk- tionsabarbeitung
Berechnend	Geht als Variable in eine Berechnung ein – beeinflusst aber nicht das Berechnungsverfahren

Abbildung 7: Wirkungsarten zur Klassifikation von Parametrisierungsauswirkungen

Um die konkreten Auswirkungen von Parametern zu erfassen und explizit darzustellen, kann man sich (in Anlehnung an [Ac04:141]) einer tabellarischen Darstellung bedienen. Dabei sind in jeder Tabellenzeile der Parametername, ein spezifikationsrelevanter Sachverhalt und dessen Kontext, eine textuelle Beschreibung und eine Wirkungsart anzugeben. Dadurch wird sowohl erfasst, worauf (Sachverhalt, Kontext) und wie (Beschreibung, Wirkungsart) sich der Parameter auswirkt. Bei den spezifikationsrelevanten Sachverhalten, auf die sich ein Parameter auswirkt, könnte es sich z. B. um eine Vorbedingung oder um ein Qualitätsprofil handeln. Zusätzlich wird der Kontext angegeben, zu dem dieser Sachverhalt gehört – beispielsweise könnte sich die Vorbedingung auf die Methode *ILagermanagement.Einlagern* oder das Qualitätsprofil auf die Schnittstelle *ILagermanagement* beziehen.

Um detailliert zu erfassen, wie ein Parameter auf den betroffenen spezifikationsrelevanten Sachverhalt wirkt, wird eine textuelle Beschreibung erstellt. Neben der ausführlichen textuellen Beschreibung ist eine Klassifikation der Auswirkungen wünschenswert, da eine solche die Übersichtlichkeit und Verständlichkeit erhöht. Dafür lässt sich das Konzept der *Wirkungsart* einsetzen. Die *Wirkungsart* klassifiziert, auf welche Art sich ein Parameter auf die Funktionalität einer Komponente auswirkt. Mögliche Ausprägungen für die Wirkungsart inklusive einer erklärenden Beschreibung werden in Abb. 7 aufgeführt. (Teile dieser Klassifikation gehen auf [DMH99:4] zurück.) Man beachte, dass in

Abb. 7 der Begriff *Ausprägung* für die Instanzen der jeweils betroffenen Sachverhalte steht.

Merkmal	Merkmalsausprägung						
Auswirkungen auf	Datensicht		Funktionssicht			Prozesssicht	
Charakter	Identifizierend	Beschreibend		Referenzierend		Steuernd	
Reichweite der Auswirkungen	Komponentenlokal			Komponentenübergreifend			
Wirkungsart	variantenbestimmend	definierend	auswählend	einschränkend	entscheidend	entscheidungsunterstützend	berechnend

Abbildung 8: Klassifikation von Parametern nach ihren Auswirkungen

Abschließend fasst Abb. 8 nochmal solche Eigenschaften von Parametern zusammen, die beschreiben, auf welche Art ein Parameter Struktur und Verhalten einer Komponente beeinflusst. Das Merkmal *Auswirkungen auf* gibt an, welche Sichten einer Komponente durch Wahl eines Parameterwertes beeinflusst werden (vgl. Abb. 2). Das Merkmal *Charakter* unterscheidet Parameter nach ihrer Rolle bei der Anpassung (vgl. Abb. 3). Das Merkmal *Reichweite der Auswirkungen* beschreibt, ob die durch einen Parameter erzeugten Auswirkungen nur die eigene Komponente betreffen oder aber sich über Komponentengrenzen hinweg erstrecken. Das Merkmal *Wirkungsart* klassifiziert, auf welche Art sich Parameter auf die Funktionalität der Komponente auswirken (vgl. Abb. 7).

8 Zusammenfassung und Ausblick

Diese Arbeit beschäftigte sich mit den Grundlagen parametrisierbarer betrieblicher Softwarekomponenten. Dazu wurden Parameter, Parametergruppen, Parametrisierungsaufgaben und -auswirkungen als die zentralen Gegenstände der Parametrisierung bestimmt, präzise definiert und deren Eigenschaften diskutiert. Außerdem wurde ein Parameter-Metamodell vorgestellt. Als Ergebnis entstand ein besseres und systematisches Verständnis der mit Parametrisierung zusammenhängenden Konzepte betrieblicher Softwarekomponenten. Interessant für die Zukunft wäre zu untersuchen, ob sich die

Ergebnisse auf nicht-komponentenorientierte betriebliche Anwendungssysteme verallgemeinern lassen und welche Rolle Parametrisierung zur Anpassung von Services spielt.

9 Literaturverzeichnis

- [Ac02] Ackermann, J.: Spezifikation des Parametrisierungsspielraums von Fachkomponenten – Erste Überlegungen. In: K. Turowski (Hrsg.): 3. Workshop Modellierung und Spezifikation von Fachkomponenten. Nürnberg 2002, S. 17-68.
- [Ac04] Ackermann, J.: Zur Beschreibung datenbasierter Parametrisierung von Softwarekomponenten. In: Turowski, K. (Hrsg.): Architekturen, Komponenten, Anwendungen – Proceedings zur 1. Verbundtagung Architekturen, Komponenten, Anwendungen (AKA 2004). LNI Bd. P-57. Augsburg 2004, S. 131-149.
- [AR00] Appelrath, H.-J.; Ritter, J.: R/3-Einführung: Methoden und Werkzeuge. Springer, Berlin, Heidelberg 2000.
- [AT06] Ackermann, J.; Turowski, K.: Zur Rolle von Parametrisierung bei der fachlichen Anpassung betrieblicher Softwarekomponenten. In: Schelp, J. et al. (Hrsg.): Integration, Informationslogistik und Architektur – Proceedings zur DW2006. LNI Band P-90. Friedrichshafen 2006, S. 341-359.
- [AT07] Ackermann, J.; Turowski, K.: On the Specification of Parameterizable Business Components. In: Draheim, D.; Weber, G. (Hrsg.): Trends in Enterprise Application Architecture (TEAA 2006). Springer-Verlag LNCS 4473. Berlin 2007, S. 25-39.
- [Bo97] Bosch, J.: Adapting Object-Oriented Components. In: Proceedings of the 2nd International Workshop on Component-Oriented Programming (WCOP'97). Turku, Finland, 1997.
- [BRS95] Becker, J.; Rosemann, M.; Schütte, R.: Grundsätze ordnungsmäßiger Modellierung. In: Wirtschaftsinformatik 37 (1995) 5, S. 435-445.
- [BS04] Becker, J.; Schütte, R.: Handelsinformationssysteme. Domänenorientierte Einführung in die Wirtschaftsinformatik. 2. Aufl. Redline Wirtschaft, Frankfurt 2004.
- [CE00] Czarnecki, K.; Eisenecker, U. W.: Generative Programming: Methods, Tools, and Applications. Addison-Wesley, Boston 2000.
- [DFH00] Disterer, G.; Fels, F.; Hausotter, A.: Taschenbuch der Wirtschaftsinformatik. Fachbuchverlag Leipzig. Leipzig 2000.
- [DMH99] Dittrich, J.; Mertens, P.; Hau, M.: Dispositionsparameter von SAP R/3-PP : Einstellungshinweise, Wirkungen, Nebenwirkungen. Vieweg, Wiesbaden 1999.
- [DW98] D'Souza, D.F.; Wills, A.C.: Objects, Components, and Frameworks with UML: The Catalysis Approach. Addison-Wesley. Reading 1998.
- [Gö97] Görk, M.: Customizing. In: P. Mertens (Hrsg.): Lexikon der Wirtschaftsinformatik. 3. Auflage. Springer-Verlag, Berlin Heidelberg 1997, S. 101-102.
- [Gr98] Griffel, F.: Componentware. Konzepte und Techniken eines Softwareparadigmas. dpunkt Verlag. Heidelberg 1998.
- [Hu94] Hufgard, A.: Adaption betriebswirtschaftlicher Softwarebibliotheken. Dissertation. Würzburg 1994.

- [JGJ97] Jacobson, I.; Griss, M.; Jonsson, P.: Software Reuse. ACM Press /Addison Wesley Longman. New York, 1997.
- [Me05] Mertens, P.; Bodendorf, F.; König, W.; Picot, A.; Schumann, M.; Hess, X.: Grundzüge der Wirtschaftsinformatik. 9. Aufl. Springer-Verlag. Berlin/Heidelberg 2005.
- [Mo00] Monson-Haefel, R.: Enterprise JavaBeans™. 2nd Edition. O'Reilly & Associates. Sebastopol, California, 2000.
- [OMG02] OMG (Hrsg.): CORBA Components Specification. Version 3.0, June 2002. URL: <http://www.omg.org/technology/documents>, Abruf am 2007-09-19.
- [OMG05] OMG (Hrsg.): Unified Modeling Language: Superstructure. Version 2.0, formal/05-07-04. URL: <http://www.omg.org/technology/documents>, Abruf am 2007-02-02.
- [Ov06] Overhage, S.: Vereinheitlichte Spezifikation von Komponenten – Grundlagen, UnSCom Spezifikationsrahmen und Anwendung. Dissertation. Augsburg 2006.
- [Pi93] Pietsch, M.: PAREUS-RM – ein Tool zur Unterstützung der Konfiguration von PPS-Parametern im SAP-System R/2. In: Wirtschaftsinformatik 35 (1993) 5, S. 434-445.
- [Pr02] Prosise, J.: Microsoft .NET – Das Entwicklerbuch. Microsoft Press, 2002.
- [Re01] Reussner, R.: Parametrisierte Verträge zur Protokolladaption bei Software-Komponenten. Logos Verlag, Berlin 2001.
- [Ri00] Ritter, J.: Prozessorientierte Konfiguration komponentenbasierter Anwendungssysteme. Dissertation. Oldenburg 2000
- [Ro99] Ropohl, G.: Allgemeine Technologie. Hanser Verlag. München/Wien 1999.
- [SAP02] SAP (Hrsg.): SAP Implementation Guide (IMG). In: Online-Dokumentation für SAP R/3 enterprise, Release 4.70. Walldorf 2002.
- [SAP97] SAP (Hrsg.): R/3-Referenz(prozeß)modell 4.0 im R/3 Business Engineer – Zielsetzung, Inhalte, Vorgehensweise. Walldorf 1997.
- [SB00] Svahnberg, M.; Bosch, J.: Issues Concerning Variability in Software Product Lines. In: van der Linden, F. (Hrsg.): Software Architectures for Product Families. Proceedings of the Third International Workshop on Software Architectures for Product Families. Springer LNCS 1951. Las Palmas de Gran Canaria 2000, S. 146-157.
- [Sc98] Schütte, R.: Grundsätze ordnungsmäßiger Referenzmodellierung. Gabler Verlag. Wiesbaden 1998.
- [SC98] Stiemerling, O.; Cremers, A. B.: Tailorable Component Architectures for CSCW-Systems. In: Proceedings of the 6th Euromicro Workshop on Parallel and Distributed Programming. IEEE Press, Madrid 1998, S. 302-308.
- [SGR98] Stickel, E.; Groffmann, H.-D.; Rau, K.-H.: Gabler Wirtschaftsinformatiklexikon. Gabler Verlag, Wiesbaden 1998.
- [Sz02] Szyperski, C.; Gruntz, D.; Murer, S.: Component Software: Beyond Object-Oriented Programming. 2. Aufl. Addison-Wesley. Harlow 2002.
- [Tu03] Turowski, K.: Fachkomponenten: Komponentenbasierte betriebliche Anwendungssysteme. Shaker Verlag, Aachen 2003.

Serviceorientierte Architekturen – Potentiale für eine nachhaltige Steigerung der Unternehmensagilität

Joachim Schelp, Stephan Aier

Institut für Wirtschaftsinformatik
Universität St. Gallen
Müller-Friedberg-Strasse 8
CH-9000 St. Gallen
{joachim.schelp | stephan.aier} @unisg.ch

Abstract: Serviceorientierte Architekturen sind in den vergangenen Jahren aus verschiedenen Gründen eingeführt worden. Analog zur vorherigen Einführung der EAI-Technologien (Enterprise Application Integration) wurde zunächst die Begründung über Wiederverwendungs- und Kostensenkungspotentiale bemüht. Angesichts der mit der Einführung serviceorientierter Architektur steigenden Komplexität der Applikationslandschaft wird diese Begründung jedoch zu einer Enttäuschung der Erwartungen führen. Serviceorientierte Architekturen bieten jedoch ein grosses Potential, zur Steigerung der betrieblichen Agilität beitragen zu können. Um dies zu erreichen, ist ein Management der Unternehmensarchitektur notwendig, das nachhaltig angelegt sein muss. Dieser Beitrag diskutiert zunächst die Probleme, die in der Begründung serviceorientierter Architekturen über Wiederverwendung und Kostensenkung liegen. Im Kontrast dazu stehen die agilitätswirksamen Potentiale. Mit Bezug auf die darin liegenden Herausforderungen wird in diesem Beitrag anhand ausgewählter Fallstudien aufgezeigt, in welcher Form sich Unternehmen darauf vorbereiten, dieses Potential langfristig zu erschliessen.

1 Einleitung

In der Praxis wurden zunächst Enterprise Application Integration (EAI) und kurz darauf auch serviceorientierte Architekturen (SOA) als Mittel angesehen, um der Komplexität der zumeist evolutionär stark gewachsenen betrieblicher Applikationslandschaften zu begegnen, was auch in der wissenschaftlichen Diskussion aufgegriffen wurde [vgl. z.B. AS04a; Li00; ST07]. Bei einer mehr von der Technik getriebenen Einführung der EAI konnten die angestrebten Ziele im Sinne einer Verringerung der Kosten z.B. durch Standardisierung der Schnittstellen und Reduktion ihrer Anzahl in Teilen erreicht werden. Die ebenfalls angestrebten Ziele zur Steigerung der Flexibilität, wie sie z.B. bei [AS04b; Ka02; Ke02] oder den Praxisbeiträgen des Sammelbands von [AS05] anklingen, wurden jedoch nicht erreicht. Die Konfrontation mit den übertriebenen Erwartungen bei den Anwenderunternehmen hat EAI als Thema in den Hintergrund treten lassen. Stattdessen werden ähnliche Erwartungen mit serviceorientierten Architekturen verknüpft. Dabei werden nicht zuletzt aus den Erfahrungen mit EAI zunächst die Wiederverwendungs- und Kostensenkungspotentiale betont. Getrieben durch Softwarehersteller, Analysten

und Berater wird in jüngster Zeit die Erwartung wieder geweckt, durch das Einbringen dieser Technologie eine bessere Anpassbarkeit der Informationssysteme an die sich ändernden Geschäftsprozesse erreichen zu können und mithin eine bessere Geschäftsprozessunterstützung.

Statt Flexibilität und genereller Abstimmung zwischen Fach- und IT-Seite wird in diesem Beitrag zunächst ein an der Agilität orientiertes Zielsystem aufgebaut, in dem die anderen Ziele positioniert werden. Da dieses Ziel nicht nur einmalig, sondern dauerhaft im Sinne eines kontinuierlichen Prozesses anzustreben ist, wird zudem hinterfragt, ob die Strukturen und Prozesse zur Zielerreichung nachhaltig gestaltet sind. Die sich ergebenden Anforderungen werden in Abschnitt 2 skizziert. Da gegenwärtig erst wenige Unternehmen entsprechende Prozesse und Strukturen etabliert haben, wird der gegenwärtige Stand der Praxis in Form von ausgewählten Fallstudien erschlossen (Abschnitt 3), die unter dem Gesichtspunkt der Nachhaltigkeit dann analysiert werden (Abschnitt 4). Fazit und Ausblick auf weitere Forschungsfragen beschließen diesen Beitrag. Die Zieldiskussion sowie die grundsätzliche Bewertung von SOA erfolgen anhand einer Literaturanalyse, die Bewertung der organisatorischen Aspekte anhand einer explorativen Fallstudienanalyse im Sinne von Yin [Yi02].

2 Ziele des Unternehmens und der betrieblichen IS

Ausgehend von der Annahme, dass die Unternehmen sich in einem intensiven, wenn nicht intensivierenden Wettbewerbsumfeld befinden, wird unterstellt, dass die Steigerung der Agilität des Unternehmens eine bedeutende Rolle in dessen Zielkatalog einnimmt. Weiterhin wird angenommen, dass dieses Ziel auch der betrieblichen IT als der für die IT-Unterstützung der Geschäftsprozesse zuständigen Organisationseinheit als wesentliches Ziel vorgegeben wird und in der Gestaltung der von ihr bereitgestellten Informationssysteme zum Ausdruck kommen muss.

2.1 Agilität

Trotz der immer wieder anzutreffenden Gleichsetzung des Begriffs Agilität mit Flexibilität, wird hier ein weitergehendes Verständnis des Agilitätsbegriffs verwendet, das auf die produktionswirtschaftliche Diskussion der vergangenen Jahre Bezüge nimmt. Z.B. findet sich bei [GNP95; SZ99; YSG99; ZS00] ein weit über Flexibilität hinausgehendes Verständnis von Agilität: Unter Flexibilität wird die Fähigkeit eines Systems verstanden, sich an erwartete Änderungen anzupassen; Agilität hingegen schliesst auch die Anpassungsfähigkeit an unerwartete Änderungen ein, was z.B. bei [Be01; VF98] eingeführt wird. In der Produktionswirtschaft wird dabei versucht, mehr „eingebaute“ Flexibilität durch Konfigurierbarkeit schon im Entwurf zu berücksichtigen – sowohl der Produktionsstrukturen, z.B. durch das Verwenden hochkonfigurierbarer CIM und CAD/CAM-Systeme, wie der Produkte selber, z.B. durch einen komponentenbasierten Entwurf. Konfigurierbarkeit fördert die Flexibilität, nützt jedoch nicht oder nur eingeschränkt bei unerwarteten Änderungen, da nur die erwarteten Änderungen berücksichtigt werden konnten. Konsequenterweise wird Agilität nicht nur für die Produktionssysteme sondern

für das ganze Unternehmen gefordert [Be01; DLP97]. Yusuf et al. definieren diese Anforderung als: „Agility is the successful exploration of competitive bases (speed, flexibility, innovation proactivity, quality and profitability) through the integration of reconfigurable resources and best practices in a knowledge-rich environment to provide customer-driven products and services in a fast changing market environment” [YSG99]. Dabei ist es wichtig festzuhalten, dass Agilität nicht auf das *Reagieren* abstellt, sondern den Wandel *pro-aktiv* unterstützen soll [GNP95]. An den fünf Unterzielen der Definition von Yusuf et al. orientiert sich die weitere Untersuchung und Bewertung des Agilitätsbeitrags serviceorientierter Architekturen in Abschnitt 2.3. Dabei soll jedoch auch berücksichtigt werden, dass Agilität nicht nur für einen Zeitpunkt erreicht sein muss, sondern dauerhaft: Es ist ein nachhaltiges Vorgehen erforderlich. Es stellt sich jedoch die Frage, wie Nachhaltigkeit erreicht werden kann.

2.2 Nachhaltigkeit

Bei der Analyse der Definitionen und Ansätze der Nachhaltigkeit ist die Herkunft dieses Begriffs aus der Umweltökonomie nicht zu übersehen. Die wohl meist verbreitete Definition stammt aus dem so genannten Brundtland Report: “Sustainable development seeks to meet the needs and aspirations of the present without compromising the ability to meet those of the future” [Wo87]. Es existiert eine beliebige Anzahl weiterer Definitionen [Co00; Te02]. Huber bündelt basierend auf der Analyse der relevanten Literatur Nachhaltigkeit in die drei Strategien Effizienz, Suffizienz und Konsistenz [Hu95], Gronau fügt die Strategie der Partizipation hinzu [Gr03].

Der *Suffizienzstrategie* liegt die Frage zu Grunde: Wie viel ist genug? Eine Antwort lässt sich nur schwer finden, auf jeden Fall scheint es aber angebracht, genügsam zu sein. Die Kritik an dieser Strategie lautet: Sie ist unrealistisch und mit Fehlwirkungen behaftet, weil sie der allgemeinen Norm der individuellen Nutzenmaximierung zuwider läuft und zu (wirtschaftlicher) Stagnation oder zumindest zu Fehlentwicklungen führen kann. Dem kann entgegen gehalten werden, dass es keine Unbeschränktheit gibt, da jedes System Grenzen in Raum und Zeit aufweist. Es sollten darum ökonomische statt moralische Anreize gesetzt werden [Hu95]. Die *Effizienzstrategie* zielt auf die Steigerung der Produktivität, um dadurch Leistungen wirtschaftlich, d.h. mit dem kleinsten möglichen Ressourcenverbrauch zu erstellen, womit zugleich einem Unterziel der Agilität entsprochen wird. Zentrale Konzepte sind dabei Wiederverwendung und Langlebigkeit. Die Effizienzstrategie ist die für das Wirtschaftsgeschehen anschlussfähigste Strategie, weshalb sie hier auch oft mit Nachhaltigkeit verwechselt wird. Die *Konsistenzstrategie* hat entweder die vollständige Abschirmung von Systemen von deren Umwelt oder die Sicherstellung deren Stimmigkeit mit dem sie umgebenden System zum Ziel [Hu95]. Die *Partizipationsstrategie* schließlich fordert die Teilhabe, d.h. die Einbindung der von einer Systemgestaltung Betroffenen. Dies ist nötig, um das System zum einen bestmöglich zu gestalten und zum anderen, um die Nutzerakzeptanz sicherzustellen [Gr03].

Wird Nachhaltigkeit als ein für die Unternehmung relevantes Thema betrachtet, so geht es meist um die Reduktion negativer externer Effekte auf die physische und soziale Umwelt [LS03]. Dies soll im Folgenden als *extern orientierte Nachhaltigkeit* bezeichnet

werden. Unser zugrunde gelegtes Verständnis von Nachhaltigkeit sei losgelöst von ökologischen Aspekten. Hahn/Hungenberg definieren das Oberziel einer jeden Unternehmung als Erhaltung und erfolgreiche Weiterentwicklung, als Erfüllung der Individualziele aller an der Unternehmung interessierten Gruppen [HH01]. Das Ziel der Betrachtung soll eine unternehmensinterne Sicht der Nachhaltigkeit mit dem primären Ziel der langfristigen, effizienten Unternehmensführung und -erhaltung sein, die auch dem Agilitätsziel Rechnung trägt. Dies kann als *intern orientierte Nachhaltigkeit* bezeichnet werden.

Aspekte der *Konsistenzstrategie* sind auf Seiten der Organisation im interdependenten Charakter der Teildimensionen der Organisationsarchitektur wieder zu finden. Es herrscht Einigkeit darüber, dass Organisationsstruktur und die darin ablaufenden Geschäftsprozesse nicht unabhängig voneinander gestaltet werden können, da es sich um unterschiedliche Sichtweisen auf den gleichen Betrachtungsgegenstand handelt. Daher würde eine mangelnde Stimmigkeit zwischen beiden Dimensionen zu Ineffizienzen führen [BI91]. Ebenso ist auf Seiten der IT eine Konsistenz zwischen den Informationssystemen notwendig. Konsequenterweise ist dann für das Gesamtsystem Unternehmung Konsistenz zwischen der Organisationsarchitektur und der Informationssystemarchitektur notwendig. Auf Seiten der IT fordert die *Effizienzstrategie* einen minimalen Anpassungsaufwand der IT. Änderungen oder Erweiterungen am System sollen schnell und kostengünstig erfolgen. Gleichzeitig soll jedoch die Komplexität des Systems so wenig wie möglich erhöht werden [Ha03]. Wird das Gesamtsystem Unternehmung verändert indem z.B. neue Technologien (bspw. Web Services) oder neue Gestaltungsparadigmen (bspw. serviceorientierte Architekturen) eingebracht werden, ergeben sich Änderungen des Gesamtsystems, die z.B. in neuen Organisationsstrukturen, Richtlinien, Führungsinstrumenten etc. münden können. Im Sinne der *Suffizienzstrategie* muss dabei gefordert werden, dass die notwendigen Änderungen ausreichend sind, aber zugleich nicht überdimensioniert werden. Um die notwendigen Veränderungen innerhalb der Organisation und der IT erfolgreich zu bewältigen, sind Akzeptanz durch Einbindung und Mitwirken aller Betroffenen an der Transformation notwendig (*Partizipationsstrategie*).

Zusammenfassend lässt sich festhalten, dass serviceorientierte Architekturen nur dann nachhaltig zur Verbesserung der Agilität eines Unternehmens beitragen, wenn Sie wie in Abschnitt 2.1 ausgeführt positive Beiträge auf die Unterziele Geschwindigkeit (der Änderungsanpassung), Flexibilität, pro-aktiver Innovation, Qualität und Profitabilität aufweisen und dabei sowohl der Effizienzstrategie (Profitabilität) genügen, als auch die Anforderungen der Konsistenzstrategie (Übereinstimmung mit Organisationsarchitektur) und der Partizipationsstrategie (Einbindung der Fachseite) ebenso berücksichtigen, wie die der Suffizienzstrategie. Bei letzterer ergibt sich die Anforderung, dass die Maßnahmen zur Sicherstellung der verschiedenen Ziele diese zwar explizit adressieren, aber dabei nicht im Sinne einer Bürokratisierung über das Ziel hinausschießen.

2.3 Eine erste Bewertung serviceorientierter Architekturen

Inwiefern kann nun eine serviceorientierte Architektur zur Erhöhung der Agilität beitragen? Entsprechend der eingangs aufgeführten Definition von Agilität nach Yusuf et al. beinhaltet Agilität die folgenden Unterziele [YSG99]: (1) Geschwindigkeit, (2) Flexibili-

tät, (3) Pro-aktive Innovation, (4) Qualität und (5) Profitabilität. Der Aspekt der pro-aktiven Innovation wird bei dieser Betrachtung außer Acht gelassen, da er primär durch die betriebliche Organisation der Informationsverarbeitung im Unternehmen beeinflusst wird, weniger durch die gewählten Architektur-Paradigmen der Informationsverarbeitung. Da serviceorientierte Architekturen mit unterschiedlichen Technologien implementiert werden können – vgl. z.B. die unterschiedlichen verwendeten Technologien in [Ha03; HBK04] – würde eine detaillierte Betrachtung der in Frage kommenden Technologien den Rahmen dieses Beitrags sprengen. Auch soll an dieser Stelle nicht in die „IT als Enabler“-Diskussion eingetreten werden.

Für den Aspekt der Qualität kann festgehalten werden, dass moderne Technologien stärker als in der Vergangenheit die Erstellung von Dokumentation erzwingen bzw. mehr Code generiert wird als bei früheren Ansätzen. Auch wird bei den neueren Infrastruktur-Generationen (im Sinne der verwendeten Middleware) der Einsatz von Repositories gefördert, so dass die manuelle Suche nach bestehenden Komponenten vermieden werden kann. Allein dieses über Repositories leichtere Auffinden vorhandener Komponenten stellt nicht selten einen Qualitätssprung in den Entwicklungsprozessen dar. Da aber auch Werkzeuge unterschiedlicher Hersteller zum Einsatz kommen, die oft nicht kompatibel miteinander sind, kann keine generelle Aussage darüber getroffen werden, ob bei einer moderneren Infrastruktur, die beim Aufbau serviceorientierter Architekturen zum Einsatz kommen kann, tatsächliche Qualitätsvorteile gegenüber früheren Ansätzen erzielt werden. Ein erhöhtes Potential dazu liegt jedoch vor.

Eine serviceorientierte Architektur kann auch insofern zur Agilität beitragen, als dass durch standardisierte Infrastruktur-Schnittstellen die verschiedenen beteiligten Systeme leichter kombiniert werden können. Durch diese „eingebaute“ Anpassbarkeit an erwartete Änderungen wird die *Flexibilität* potentiell erhöht. Um diese Vorteile in vollem Umfang nutzen zu können, muss jedoch auch eine System- und Schnittstellen-Standardisierung erfolgen. Aber auch wenn Standardsysteme nicht verfügbar sind und Eigenentwicklungen die Applikationslandschaft dominieren, bietet eine serviceorientierte Architektur im Sinne einer höheren Flexibilität Vorteile. Die Standardisierung wird erleichtert, da auf der Infrastrukturebene zahlreiche Komponenten angeboten werden, auf deren Basis Eigenentwicklungen mit technisch standardisierten Schnittstellen vorhanden sind.

Sie bietet aber vor allem im Sinne des Ziels der *Geschwindigkeit* Vorteile, indem sie eine Verkürzung der Entwicklungszeiten (fachlicher) Komponenten eröffnet. Und dies nicht nur bei der Entwicklung neuer Komponenten, sondern auch bei der Anpassung bestehender Systeme. Dieser Vorteil wird anhand des folgenden Szenarios erläutert: Es sei angenommen, dass eine fachliche Anforderung nach IT-Unterstützung zur Entwicklung eines fachlichen Services geführt hat, der seinerseits aus teilweise vorhandenen, teilweise neu entwickelten Komponenten (atomaren Services) zusammengesetzt wird. Die Entwicklung dieses Services hat durch die Wiederverwendung vorhandener Bausteine beschleunigt erfolgen können. Es sei weiterhin angenommen, dass die Fachseite im Zeitablauf eine Änderung an diesem Service wünscht. Diese Änderung muss nun je nach Ausmass in unterschiedlicher Form umgesetzt werden: Im einfachsten Fall handelt es sich um eine Änderung, die – im Sinne einer „black box“ – innerhalb des Services vor-

genommen werden kann und die Schnittstellen des Services nicht verändert. In diesem Fall sind keine Anpassungen in von diesem Service abhängigen Komponenten notwendig. Dabei sei angenommen, dass sich der Testaufwand bei dieser Änderung ebenfalls in Grenzen hält, da unterstellt wird, dass allfällige Tests nur gegen die mit dem Service verbundenen Komponenten durchgeführt werden.

Wenn die gewünschte Änderung jedoch zur Konsequenz hat, dass eine Veränderung der Schnittstelle(n) eines der beteiligten Services notwendig ist, erhöht sich der Änderungsaufwand. In einer nicht-serviceorientierten (bzw. nicht-komponentenorientierten) Architektur müssen alle abhängigen Komponenten an die veränderte Schnittstelle angepasst werden. In einer Serviceorientierten Architektur hingegen, in der die Vorteile der Objekt- und Komponentenorientierung zum Tragen kommen, kann die gleiche Schnittstelle in unterschiedlichen Varianten angeboten werden, die sich durch ihre Signatur (Parameterkombination) unterscheiden. Die von der alten Variante abhängigen Services müssen nicht angepasst werden, für diese verhält sich die geänderte Komponente wie zuvor. Der Geschwindigkeitsgewinn durch nicht notwendige Abhängigkeitsänderungen liegt auf der Hand. Später hinzugefügte Services, die auf die neue Funktionalität zugreifen müssen, können dagegen über die neue Variante darauf zugreifen. Im einfachsten Fall kann die zusätzliche Variante über einen zusätzlichen Regelsatz innerhalb der Integrationsinfrastruktur (Integrationsbus) abgebildet werden. Sofern keine Logik auf dem Bus zulässig ist oder zu komplex für diesen ist, muss eine explizite Variante erstellt werden. In der Praxis kommt dieser Fall durchaus vor, es wird in Einzelfällen von bis zu sechs produktiven Varianten eines Service berichtet.

Die nicht zuletzt durch die Varianten steigende Zahl der Services ist jedoch problematisch. Jede zusätzliche Variante eines Service verringert das Wiederverwendungspotential des zugrundeliegenden Services. In der Praxis zeigt sich, dass die Wiederverwendung nur bei wenigen Services sehr hoch ist, der durchschnittliche Wiederverwendungsgrad hingegen sehr niedrig. Ein eindrucksvolles Beispiel hierzu findet sich z.B. bei [SH06], in dem bei einer serviceorientierten Architektur mit zum Erhebungszeitpunkt 650 Services zwar ein prozentualer Anteil wiederverwendeter Services von 34%, aber nur eine durchschnittliche Wiederverwendung von 1.7 konstatiert wurde. Bemerkenswert ist dabei insbesondere, dass sich die Fälle hoher Wiederverwendung auf nur sehr wenige Services konzentrieren (14 Services wurden 10 mal oder mehr wiederverwendet; zu den dennoch sehr positiven Erfahrungen mit dieser Architektur im skizzierten Fall vgl. insbesondere [Ha03]). Mit jeder Variante steigt zudem der Wartungsaufwand. Daher muss eine serviceorientierte Architektur hinsichtlich der *Profitabilität* sehr vorsichtig beurteilt werden. Insbesondere wenn der monetäre Nutzen eines SOA-Projekts nicht bestimmt werden kann. Ein weiterer Nachteil ist die steigende Komplexität des Gesamtsystems, die sich indirekt und eher langfristig negativ auf Qualität, Geschwindigkeit, Flexibilität, Innovationsfähigkeit und auch Profitabilität auswirken kann. Unterschiedliche Varianten eines Services, die jeweils in unterschiedlicher Anzahl Wiederverwendung finden, können langfristig zu einem sehr eng vermaschten Netz führen, dessen Komplexität größer ist als die des mit der serviceorientierten Architektur abgelösten Monolithen. Damit besteht in langfristig die Gefahr, dass durch die bei einer SOA stärker steigende Komplexität – aufgrund einer im Vergleich zu Legacy-Systemen stärker steigenden Zahl von Abhängigkeitsbeziehungen – die Agilität negativ beeinflusst werden kann. Kurzfristig ist

eine höhere Agilität erreichbar, da aufgrund der Varianten schneller reagiert werden kann.

Zusammenfassend kann festgehalten werden, dass Indizien dafür bestehen, dass eine serviceorientierte Architektur in der kurzen Frist zum Erreichen einer höheren Agilität beitragen kann. Es stellt sich jedoch die Frage, inwiefern dies auch nachhaltig ist: Einerseits kann in der langen Frist aufgrund der höheren Komplexität die Flexibilität sinken, andererseits hängt die Bewertung auch von der Gewichtung der Unterziele der Agilität ab. Da dies noch stärker von unternehmensspezifischen Faktoren abhängt als der unmittelbare Beitrag zur Agilität, erfolgt im weiteren Verlauf eine explorative Fallstudienanalyse, aus der sich Ansatzpunkte für weitere Forschungsarbeiten ergeben.

3 Fallbeschreibungen

Nachfolgend werden fünf Unternehmen skizziert, die in den letzten Jahren eine serviceorientierte Architektur aufgebaut haben. Es wird neben einer groben Beschreibung des Unternehmens kurz dargelegt, was die Motivation zur Einführung der serviceorientierten Architektur war, welche Architekturebenen unterschieden werden und in welcher Form Gestaltung und Betrieb dieses Architekturkonzepts erfolgen. Hinsichtlich der Architekturebenen erfolgt zur leichteren Vergleichbarkeit eine Übersetzung auf die in [WF06] definierten Geschäfts-, Organisations- (Ablauf- und Aufbau), Integrations-, Software- und Infrastrukturebene. Bei den Architekturmanagementprozessen wird betrachtet, inwieweit eine Abstimmung zwischen den Ebenen erfolgt und in welchem Umfang die Fachseite beteiligt ist.

3.1 Unternehmen A

Unternehmen A ist eine der größten Banken der Schweiz. Unternehmenszusammenschlüsse in der Historie des Unternehmens hatten langfristig Folgen auf die Komplexität der entstandenen Applikationslandschaft. Mit dem steigenden Integrationsbedarf erwuchs die Motivation zur Einführung einer serviceorientierten Architektur. Im Jahr 2002 bestand das Kernbankensystem aus über 450 Host-basierten sowie Client-Server-Systemen. Um der daraus resultierenden zunehmenden Integrationskomplexität zu begegnen, wurde im Jahr 2001 eine SOA-Vision entwickelt. Erste Schritte zur Umsetzung bestanden in der Kapselung der Funktionalität bestehender Systeme als fachliche Services bzw. in der direkten Implementierung neuer Funktionalität als fachliche Services. Die Kapselung erbrachte Verbesserungen hinsichtlich Flexibilität, indem Integrationsbedarfe besser umgesetzt werden konnten und Verbesserungen hinsichtlich der Geschwindigkeit, da ausgewählte Komponenten zunehmend besser wiederverwendet und schnelle Projektdurchführungszeiten erzielt werden konnten. Hinsichtlich der aufgebauten Architektur werden eine Geschäftsarchitektur, eine Applikationsarchitektur, eine Integrationsarchitektur, eine Softwarearchitektur, eine Komponentenarchitektur sowie eine technische Architektur unterschieden. Die IT-bezogenen Architekturen sind dabei jedoch wesentlich stärker ausgeprägt als die Geschäftsarchitektur: Die Geschäftsmodelle sind nur in geringem Umfang explizit modelliert, bei den Geschäftsprozessen sind Umfang

und Aktualität je nach Fachbereich unterschiedlich. Die Architekturmanagementprozesse sind in Bezug auf die betriebliche IT stark ausgeprägt: Ein Team von insgesamt 90 Architekten sorgt für eine klare Architekturkommunikation und -durchsetzung in Richtung Entwicklung. Hierfür existieren klar definierte Architekturprozesse, die eine starke IT-Governance widerspiegeln. In Richtung der Fachseite ist die Wirkung jedoch weniger stark ausgeprägt. Zwar gibt es eine starke Architekturposition in Bezug auf einzelne IT-Projekte. Den umfangreichen Modellierungsanstrengungen auf IT-Seite stehen jedoch auf der Fachseite keine äquivalenten Strukturen gegenüber, hier sind Modelle und Modellierung stark projektgetrieben.

3.2 Unternehmen B

Unternehmen B ist einer der größten Energieversorger Deutschlands. Während EAI als eher technischer Ansatz verstanden und betrieben wurde, hat man beim Thema SOA frühzeitig begonnen, eine SOA-Governance zu adressieren. Neben technischen Detaillösungen wurden so – zuerst getrieben aus der Konzern-IT – Governance-Modelle entwickelt. Es wurden frühzeitig die Business-Owner eingebunden, so dass heute tatsächlich schneller neue Geschäftsprozesse technisch – auf Basis einer SOA – umgesetzt werden können und ein Beitrag zu einer erhöhten Agilität geleistet wurde. Die in diesem Fall IT-getriebene SOA blieb jedoch auf ausgewählte Bereiche der Fachgebiete beschränkt. Der Grund mag darin liegen, dass es keinen umfassenden Enterprise Architecture Ansatz gibt, wodurch beispielsweise die IT vornehmlich IT-Architekturmanagement betreibt und dies nur in ausgewählten Bereichen mit einem Geschäftsprozessmanagement kombiniert. Die darin sehr grob verankerten Architekturebenen sind die Prozessebene welche die ablauforientierten Elemente der Organisationsebene umfasst, die Architektur- und Business-Service-Ebene welche Elemente der Integrationsebene umfasst und die Ebene der Basic-Services, welche Elemente der Softwareebene beschreibt. Ebenso wie bei Unternehmen A, sind vor allem die technisch orientierten Ebenen ausgebaut, während Elemente der Organisationsebene nur rudimentär betrachtet werden.

3.3 Unternehmen C

Unternehmen C ist ein großer Finanzdienstleister in der Schweiz, der sich primär auf standardisiertes Retail-Banking und Transaktionsabwicklung fokussiert. Die im Zeitablauf entstandene Applikationslandschaft mit ihren komplexen Abhängigkeitsbeziehungen führte bei zunehmendem Integrationsbedarf zunächst zu einem umfangreichen EAI-Projekt. Daraus erwuchs dann eine SOA-Vision, um die auf IT-Seite realisierten Vorteile hinsichtlich beschleunigter Projektdurchführung, erzielter Wiederverwendung und daraus resultierender Kostensenkungen auch auf der Fachseite zu realisieren. Hinsichtlich der betrachteten Architekturmodelle finden sich in diesem Unternehmen Ausprägungen auf allen genannten Ebenen wieder. Auf Seite der betrieblichen IT bestehen umfangreiche, definierte Architekturprozesse, mit denen die serviceorientierte Architektur durchgesetzt und weiterentwickelt wird. Ursprünglich war auf Seite der betrieblichen IT die Absicht vorhanden, aus der bei ihr angesiedelten Architektur heraus auch die fachlichen Architekturen expliziter in die bestehenden Architekturmanagementprozesse einzubinden. Dies ist jedoch zugunsten einer explizit auf der Fachseite angesiedelten Organisati-

onseinheit für Gestaltung und Betrieb der Unternehmensarchitektur aufgegeben worden. Die Abstimmung zwischen den Organisationseinheiten für Unternehmens- und IT-Architekturen erfolgt explizit, wobei die personelle Verflechtung durch Einbindung (ehemaliger) IT-Architekten auf der Fachseitenorganisation die Abstimmungsprozesse erleichtert. Die über die verbesserte, letztlich serviceorientierte Applikationslandschaft erzielten Verbesserungen hinsichtlich Wiederverwendung und Kostensenkungen führten zu einer Verbesserung der Agilitätsposition. Der Beitrag der besseren Abstimmung zwischen Fach- und IT-Seite kann momentan hinsichtlich der Agilitätswirkung noch nicht abschliessend beurteilt werden, da die entsprechenden Strukturen noch nicht lange genug wirken. Erste Beobachtungen deuten aber auf einen weiterhin positiven Beitrag hin.

3.4 Unternehmen D

Unternehmen D ist ein großes, in Deutschland tätiges Telekommunikationsunternehmen. Die Telekommunikationsbranche ist im hier betrachteten Kontext von zwei wesentlichen Merkmalen bestimmt. Erstens handelt es sich um eine grundsätzlich technikaffine und technikgetriebene Branche; zweitens ist die Schnelligkeit der Umsetzung neuer Produktideen (neue Preismodelle, technisch neue Angebote) sehr hoch, da dies eine der wenigen Differenzierungsmöglichkeiten gegenüber den Wettbewerbern darstellt. Aus diesen beiden Gründen heraus, wurde frühzeitig ein Enterprise Architecture Projekt gestartet, welches einen konkreten Rahmen für technische Veränderungsprojekte bereitstellt. Dieser Rahmen ist zum einen hilfreich, um die Auswirkungen von geplanten Veränderungen schnell identifizieren zu können und um die Konformität der geplanten Änderungen bezogen auf definierte Architekturregeln prüfen zu können. Insbesondere wurden jedoch auch Prozesse definiert, die es ermöglichen, temporär bestimmte Architekturregeln unter der Maßgabe zu verletzen, dass sowohl ein Projektplan, als auch ein Projektbudget für die Wiederherstellung der Architekturkonformität existiert. So ist es möglich, weitere Geschwindigkeitsvorteile zu erzielen und positiv hinsichtlich der Agilität zu wirken. Grundsätzlich sind in einem solchermaßen technologiegetriebenen Unternehmen die Geschäftsbereiche und die IT nie sehr weit voneinander entfernt. Um dies auch dauerhaft sicherzustellen, müssen alle Veränderungsprojekte einen klaren Business Case vorweisen. In einem so stark durch den täglichen Veränderungsdruck getriebenen Unternehmen ist es auch möglich, solch einen Business Case für SOA-Projekte bzw. eine SOA-Infrastruktur zu rechtfertigen. Der Erfolg dieser Vorgehensweise liegt jedoch in den klaren Regeln und Vorgaben (Governance), welche durch das Enterprise Architecture Management vorgegeben wurden. Das zugrunde liegende Architekturverständnis hat Entsprechungen auf allen genannten Ebenen und manifestiert sich in einem für alle Veränderungen maßgeblichen Domänenmodell.

3.5 Unternehmen E

Unternehmen E ist ein weiteres Telekommunikationsunternehmen, dass unter anderem in Deutschland tätig ist, deren Konzernmutter jedoch im nicht-deutschsprachigen Raum residiert. Wie bei D sind Unternehmen und die Fachabteilungen technikaffin und das Geschäft technikgetrieben. Ebenso ist aufgrund des hohen Marktdrucks Schnelligkeit ein dominierendes Ziel, das Entwicklung, Einführung und Inbetriebnahme neuer Geschäfts-

modelle und diese unterstützender Anwendungen prägt. Über die Jahre ist daher wie bei D die Applikationslandschaft schnell gewachsen und hat eine hohe Komplexität erreicht. In der Vergangenheit wurden die Applikationen zunächst landesspezifisch entwickelt und eingeführt, zu einem späteren Zeitpunkt setzt eine Zentralisierung ein, die auch zu einer länderübergreifenden Abstimmung der zuvor verteilten und uneinheitlichen Architektur-Aktivitäten führte. Aufgrund eines Wechsels der Unternehmensstrategie ist diese länderübergreifende Abstimmung jedoch wieder zurückgenommen worden, Serviceorientierung steht bei der Entwicklung neuer Applikationen im Vordergrund, um die weiter wachsende Komplexität der (verteilten) Applikationslandschaft besser beherrschen zu können. Das Serviceverständnis ist mit dem skizzierten kompatibel. Jedoch sind die Prozesse im Sinne einer Prozessarchitektur nicht durchgängig erfasst, so dass die Abstimmung mit der Fachseite an dieser Stelle erschwert wird. Da die Architekturentwicklung konzernweit weniger stark zentral koordiniert, sondern nur in schwächerer Form abgestimmt wird, kann eher von einer föderal strukturierten Gesamtapplikationslandschaft gesprochen werden. Innerhalb der einzelnen Landesgesellschaften sind Vorgaben, Richtlinien und Governance-Strukturen unterschiedlich ausgeprägt. Mit Blick auf die deutsche Landesgesellschaft kann einer hohen Architekturkompetenz und ein vertieftes Verständnis serviceorientierter Architekturen festgestellt werden. Hier sind auch klarere Regeln und Vorgaben im Sinne einer Governance gegeben. Durch die sich ändernde Geschäftsstrategie und deren organisatorischen Auswirkungen in kürzeren Zyklen kann jedoch nur schwer eingeschätzt werden, in welchem Ausmass die etablierten Regeln, Prozesse und Strukturen auch langfristig wirken können.

4 Fallanalyse – Beitrag zur nachhaltigen Sicherung der Agilität

Vor dem Hintergrund der beschriebenen Fälle kann nun hinterfragt werden, inwieweit die serviceorientierten Architekturen und zugehörigen Managementprozesse auch nachhaltig zu einer Erhöhung der Agilität beitragen können, wobei der kurzfristige Beitrag der serviceorientierten Architekturen schon als positiv eingeschätzt werden kann. Bezogen auf die *Konsistenzstrategie* kann festgehalten werden, dass bei Unternehmen A gegenwärtig die Nachhaltigkeit skeptisch beurteilt werden muss. Die Organisationsarchitektur im Sinne der Prozessarchitektur wird nicht in dem Umfang aktiv bewirtschaftet, wie es auf Seiten der Integrations-, Software- und Infrastrukturebene der Fall ist. Bezogen auf die Fachsichten werden diese überwiegend im Rahmen von Projekten eingeholt. Das Ungleichgewicht zwischen beiden Seiten ist erkannt und erste Maßnahmen werden ergriffen, führen aber kurzfristig zu keiner vollständigen Lösung des Konsistenzproblems. Eine ähnliche Aussage gilt für Unternehmen B. Hier wurden zwar fast mustergültig eine Governance und die Schnittstellen zum Geschäftsprozessmanagement für SOA-Projekte definiert, doch beschränken sich diese Aktivitäten auf singuläre Punkte, da der stabilisierende Rahmen eines ausgewogenen Architekturmanagements nicht vorhanden ist. Bei Unternehmen C hingegen kann festgestellt werden, dass die Bedeutung einer konsistenten Modellierung über die verschiedenen Ebenen hinweg nicht nur erkannt, sondern auch in Form expliziter Architekturteams adressiert ist. Zugleich werden die Gestaltungsprinzipien serviceorientierter Architekturen konsistent angewendet. Gleiches kann für Unternehmen D konstatiert werden: Die kulturelle Nähe der Fachabteilungen zur IT sowie das vollständig und dennoch pragmatisch ausgeprägte Architekturma-

nagement schaffen eine ausgezeichnete Basis für die Verbreitung serviceorientierter Architekturen. Analog verhält es sich bei Unternehmen E (im Fokus auf Deutschland), wenngleich wie bei D die zur IT-Architektur äquivalente Geschäfts- bzw. Prozessarchitektur als Einschränkung angesehen werden kann.

Ähnlich verhält es sich bei der *Partizipationsstrategie*. Bei Unternehmen A sind die Vertreter der Fachseite nur eingeschränkt involviert. Im Rahmen der IT-Projekte geben sie zwar den Anstoß, sind aber über die Projekte hinaus nicht an der Weiterentwicklung der Architekturen bzw. dem Abgleich zwischen Fach- und IT-Architektursichten involviert. Bei Unternehmen B ist die Partizipation durch ein von der IT getriebenes SOA-Governance-Modell grundsätzlich vorbildhaft, jedoch bleibt die Partizipation auf ausgewählte, technikaffine Bereichen beschränkt. Bei Unternehmen C hingegen ist die Beteiligung der Fachseite durch Ansiedlung der für die Unternehmensarchitektur zuständigen Organisationseinheit auf der Fachseite auch strukturell fest verankert worden. Hier ist das Potential für einen Beitrag zur Nachhaltigkeit der getroffenen Lösung offenkundig. Gleiches lässt sich für die Unternehmen D und E sagen, wo es eine etablierte Zusammenarbeit von Fachbereichen, IT und Architektur gibt. Dies wird unterstützt durch weit in die Fachbereiche wirkende Fragen der Technologie, welche mit den angebotenen Produkten verknüpft sind, wie auch durch die weit in die IT reichende, konsequente Anwendung von Business Cases.

Hinsichtlich der *Suffizienzstrategie* kann festgehalten werden, dass bei Unternehmen A die getroffenen Maßnahmen für die Umsetzung des Architekturmanagements umfangreich ausgefallen sind. Für die Sicherstellung der Konsistenz der IT-bezogenen Architektursichten ist dies zwar hilfreich, doch muss hier kritisch gefragt werden, ob durch die Größe der aufgebauten Organisationseinheit und der für ihren Betrieb aufgesetzten Prozesse nicht die Gefahr der Überbürokratisierung besteht und damit in der langen Frist nicht Konsistenz und Abstimmung mit den Fachbereichen gefährdet werden. Bei Unternehmen B ist der Ausbau der Unternehmensarchitektur offensichtlich nicht suffizient, um die Serviceorientierung über Pilotprojekte hinweg im Gesamtunternehmen zu verankern. Bei den Unternehmen D und E (mit Fokus auf Deutschland) ist gewissermaßen die höchste Suffizienz zu verzeichnen, da auch SOA-Projekte einen klaren und nachweisbaren Geschäftsnutzen realisieren müssen, der sich letztlich in der Generierung zusätzlichen Umsatzpotenzials widerspiegelt. Reine SOA-Projekte als Architekturmaßnahme finden hier nicht statt.

Für die *Effizienzstrategie* kann in Anlehnung an Abschnitt 2.2 zunächst von einem positiven Beitrag serviceorientierter Architekturen ausgegangen werden. Bestätigt wird dieser durch die Erfahrungen bei Unternehmen B, C, D und E. Insbesondere bei Unternehmen C sind durch Wiederverwendung Effizienzpotentiale realisiert worden. Die Entkopplung der Systeme hat zwar zu einer gestiegenen Komplexität des Gesamtsystems beigetragen, jedoch sind die Systeme nun flexibler konfigurierbar und koppelbar. Bei den Unternehmen D und E müssen SOA-Projekte zwangsläufig zu effizienten Lösungen führen, da die Fokussierung auf Business Cases anderenfalls ihre Implementierung verhindern würde.

5 Zusammenfassung und Ausblick

Wie dargelegt wurde, bieten serviceorientierte Architekturen durchaus das Potential, einen nachhaltigen Beitrag zur Steigerung der Agilität eines Unternehmens leisten zu können. Jedoch sind mit diesem Gestaltungsparadigma nicht durchgängig die Wiederverwendungs- und Kostensenkungspotentiale verbunden, die von Herstellern und Beratern häufig angeführt werden. Dafür eröffnet sich den Anwenderunternehmen aber die Chance, in IT-Projekten schneller auf (Änderungs-) Anforderungen der Fachseite reagieren zu können und so positiv zur Agilität des Unternehmens beitragen zu können. Dies wird verstärkt durch die Potentiale einer höheren Flexibilität und Nutzung schon vorhandener Komponenten, wenngleich dies wie ausgeführt eher in einer langfristigen Perspektive berücksichtigt werden kann. So positiv der Beitrag zur Agilität bewertet werden kann, so schwierig ist die Kommunikationsaufgabe für die IT-Abteilungen: Kostenreduktion durch Wiederverwendung lässt sich leichter kommunizieren und dafür notwendige Investitionen in neue IT-Infrastrukturen lassen sich entsprechend leichter begründen und durchsetzen. Positive Beiträge zur Agilität und dafür notwendige Investitionen müssen hingegen auf einen expliziten Bedarf der Fachseite stossen, der in der Forderung nach verbesserter Geschwindigkeit (*time-to-market*), Qualität, Innovationsfähigkeit oder Flexibilität mündet und mit der notwendigen Investitionsbereitschaft versehen ist.

Aus Forschungsperspektive eröffnen serviceorientierte Architekturen durchaus weiteren Forschungsbedarf über die in der Informatik geführte Diskussion der technischen Grundlagen hinaus. Die aufgezeigten Fälle belegen, dass die Unternehmen verstärkt Aktivitäten zum Management der Unternehmensarchitektur entfalten, um der mit der Einführung serviceorientierter Architekturen kräftig steigenden Komplexität der Applikationslandschaft zu begegnen: Diese Anstrengungen konzentrierten sich zwar überwiegend auf die IT-Architekturen, doch ist eine Ausweitung auf die gesamte Breite der Unternehmensarchitekturen in Ansätzen erkennbar. Wie in einer komplementären Untersuchung durch [WS08] aufgezeigt, ist parallel dazu das Entstehen geeigneter Governance-Strukturen erkennbar, welche zur Steuerung der Architekturentwicklung notwendig sind. Inwiefern die Fachseite dabei aktiv eingebunden werden kann, ist nach wie vor ein offenes Problem, dass auch aus wissenschaftlicher Sicht zahlreiche Herausforderungen birgt. Jenseits der technischen Fragestellungen bedarf es insbesondere für das Thema der Serviceorientierung der Konstruktion geeigneter Methoden, welche die notwendigen Abstimmungsprozesse zwischen Fachseite und betrieblicher IT organisatorisch so verankern, dass ein nachhaltiges Alignment zwischen diesen erfolgt.

6 Literaturverzeichnis

- [AS04a] Aier, S.; Schönherr, M. (Hrsg.): Enterprise Application Integration – Serviceorientierung und nachhaltige Architekturen. Gito, Berlin 2004.
- [AS04b] Aier, S.; Schönherr, M.: Flexibilisierung von Organisations- und IT-Architekturen durch EAI. In: Aier, S.; Schönherr, M. (Hrsg.): Enterprise Application Integration – Flexibilisierung komplexer Unternehmensarchitekturen. Gito, Berlin 2004, S. 1-59.

- [AS05] Aier, S.; Schönherr, M.: EAI als integrierendes Element einer nachhaltigen Unternehmensarchitektur. In: Aier, S.; Schönherr, M. (Hrsg.): Unternehmensarchitekturen und Systemintegration. Gito, Berlin 2005, S. 4-56.
- [Be01] Becker, F.: Organisational agility and the knowledge infrastructure. In: Journal of Corporate Real Estate 3 (2001) 1, S. 28-37.
- [BI91] Bleicher, K.: Organisation: Strategien, Strukturen, Kulturen. 2. edition, Gabler, Wiesbaden 1991.
- [Co00] Conrad, J.: Nachhaltige Entwicklung – einige begriffliche Präzisierungen oder der heroische Versuch einen Pudding an die Wand zu nageln. http://www.fu-berlin.de/ffu/download/rep_00-07.PDF, Abruf am 09.08.2007.
- [DLP97] Duguay, C. R.; Landry, S.; Pasin, F.: From mass production to flexible/agile production. In: International Journal of Operations & Production Management 17 (1997) 12, S. 1183-1195.
- [GNP95] Goldman, S. L.; Nagel, R. N.; Preiss, K.: Agile competitors and virtual organizations: strategies for enriching the customer. Van Nostrand Reinhold, New York, NY 1995.
- [Gr03] Gronau, N.: Wandlungsfähige Informationssystemarchitekturen: Nachhaltigkeit bei organisatorischem Wandel. Gito, Berlin 2003.
- [Ha03] Hagen, C.: Integrationsarchitektur der Credit Suisse. In: Aier, S.; Schönherr, M. (Hrsg.): Enterprise Application Integration – Flexibilisierung komplexer Unternehmensarchitekturen. GITO-Verlag, Berlin 2003, S. 61-81.
- [HBK04] Herr, M.; Bath, U.; Koschel, A.: Implementation of a Service Oriented Architecture at Deutsche Post MAIL. In: Proceedings, European Conference on Web Services (ECOWS 2004), Erfurt, Germany 2004, S. 227-238.
- [HH01] Hahn, D.; Hungenberg, H.: PuK – Wertorientierte Controllingkonzepte. 6. Auflage edition, Gabler, Wiesbaden 2001.
- [Hu95] Huber, J.: Nachhaltige Entwicklung durch Suffizienz, Effizienz und Konsistenz. In: Fritz, P.; Huber, J.; Levi, H. W. (Hrsg.): Nachhaltigkeit in naturwissenschaftlicher und sozialwissenschaftlicher Perspektive. Hirzel, Wissenschaftliche Verlagsgesellschaft, Stuttgart 1995, S. 31-46.
- [Ka02] Kaib, M.: Enterprise Application Integration – Grundlagen, Integrationsprodukte, Anwendungsbeispiele. DUV, Wiesbaden 2002.
- [Ke02] Keller, W.: Enterprise Application Integration. dpunkt.verlag, Heidelberg 2002.
- [Li00] Linthicum, D. S.: Enterprise Application Integration. AWL Direct Sales, Reading, Massachusetts 2000.
- [LS03] Leitschuh-Fecht, H.; Steger, U.: Wie wird Nachhaltigkeit für Unternehmen attraktiv? – Business Case für nachhaltige Unternehmensentwicklung. In: Linne, G.; Schwarz, M. (Hrsg.): Handbuch Nachhaltige Entwicklung. Leske + Budrich, Opladen 2003, S. 257-266.
- [SH06] Schwinn, A.; Hagen, C.: Measured Integration – Metriken für die Integrationsarchitektur. In: Schelp, J.; Winter, R. (Hrsg.): Integrationsmanagement. Springer, Berlin et al. 2006, S. 267-292.
- [ST07] Starke, G.; Tilkov, S. (Hrsg.): SOA-Expertenwissen – Methoden, Konzepte und Praxis serviceorientierter Architekturen. dpunkt, Heidelberg 2007.

- [SZ99] Sharifi, H.; Zhang, Z.: A methodology for achieving agility in manufacturing organisations: An introduction. In: *International Journal of Production Economics* 62 (1999) 1-2, S. 7-22.
- [Te02] Teichert, V.: *Indikatoren zur Lokalen Agenda 21: ein Modellprojekt in sechzehn Kommunen*. Leske + Budrich, Opladen 2002.
- [VF98] Vokurka, R. J.; Fliedner, G.: The journey toward agility. In: *Industrial Management & Data Systems* 98 (1998) 4, S. 165-171.
- [WF06] Winter, R.; Fischer, R.: Essential Layers, Artifacts, and Dependencies of Enterprise Architecture. In: *Proceedings, EDOCW'06 Workshop on Trends in Enterprise Architecture Research (TEAR 2006)*, Los Alamitos, CA, USA 2006, S. 30-37.
- [Wo87] World Commission on Environment and Development: *Our common future*. Oxford University Press, Oxford, New York 1987.
- [WS08] Winter, R.; Schelp, J.: Enterprise Architecture Governance: The Need for a Business-to-IT Approach. In: *Proceedings, Proceedings of the 23rd Annual ACM Symposium on Applied Computing (SAC2008)*, Mar 16-20, 2008, Fortaleza, Ceará, Brazil, New York, NY, USA 2008, S. 548-552.
- [Yi02] Yin, R. K.: *Case Study Research. Design and Methods*. 3 edition, Sage Publications, London 2002.
- [YSG99] Yusuf, Y. Y.; Sarhadi, M.; Gunasekaran, A.: Agile Manufacturing: The Drivers, Concepts and Attributes. In: *International Journal of Production Economics* 62 (1999) 1-2, S. 33-43.
- [ZS00] Zhang, Z.; Sharifi, H.: A methodology for achieving agility in manufacturing organisations. In: *International Journal of Operations & Production Management* 20 (2000) 4, S. 496-512.

SOA Implementierung in der Luftfahrtindustrie am Beispiel der Lufthansa

Rolf Kubli, Michael Burkard, Elina Rantakallio, Annette Kreczy

EDS Information Business GmbH

Schaffhauserstrasse 550

CH-8052 Zürich

{rolf.kubli | michael.burkard | elina.rantakallio | annette.kreczy}@eds.com

Abstract: Dieser Beitrag zeigt die Lösung und die Erfahrungen eines wegweisenden SOA Projektes der Luftfahrtindustrie auf. Ausgehend von der EDS „Agile Enterprise Plattform“ und „Designed for RunTM“ SOA Strategie, wird eine Implementierung der „EDS Airline SOA Plattform“ für die einheitliche Einbindung (Middleware, Security) der Lufthansa Kernapplikationen in die „Common IT Plattform“ der Star Alliance dargestellt. Die Diskussion der Herausforderungen und Erfahrungen der Projektentwicklung, Einführung und Produktion bestätigt den erwarteten Nutzen eines SOA-Ansatzes, zeigt aber auch die Bedeutung des technischen Designs, der Testplanung, der Managementlösung und der Governance auf.

1 Einleitung

Ein Passagier der Austrian Airlines ruft im Call Center der Lufthansa an und möchte sein Business Ticket Wien über Frankfurt nach New York auf First Class ändern (upgrade) und mit seinem Meilenguthaben bezahlen.

Können Sie sich vorstellen, was ein solches Szenario-Beispiel für die IT-Systeme der betroffenen Fluggesellschaften bedeutet? Nun, es stellt grosse Anforderungen an die Integration aller wichtigen Applikationen (Reservation, Inventar, Check-in, Ticketing und Vielflieger Programm) der verschiedenen eigenständigen, aber als Mitglieder der Star Alliance zusammenarbeitenden Unternehmen. Es geht dabei darum, Transaktionen über unterschiedliche Datenformate, Protokolle und Generationen von Systemen schnell und zuverlässig abzuwickeln. Dies lässt erahnen, dass eine Service-orientierte Architektur ein vielversprechender Lösungsansatz für ein grundsätzliches, typisches Problem der Zusammenarbeit in der Transportbranche sein könnte und warum das Thema auf immer grösseres Interesse bei vielen Fluggesellschaften stösst.

Nach der häufig zitierten Definition von OASIS ist eine Service-orientierte Architektur (SOA) ein Paradigma für die Strukturierung und Nutzung verteilter Funktionalität, die von unterschiedlichen Besitzern verantwortet wird [Wi06]. Es gibt jedoch eine Reihe anderer Beschreibungen für eine SOA, so etwa auch die folgende. Eine Service-orientierte Architektur (SOA) meint ein allgemeines Modell der Datenverarbeitung geprägt von lose gekoppelten Elementen (Dienste, Funktionen, Komponenten)

und als Service exponierten Applikationen („wrapped“ legacy systems), welche über ein intelligentes Netzwerk publiziert, konsumiert und dynamisch kombiniert werden.

2 Herausforderungen in der Transportindustrie

Die Unternehmen der Transportindustrie sehen sich einer Reihe von Herausforderungen gegenüber, von der Globalisierung, über neue Konkurrenten, z.B. Billigflieger, bis zu ständig ändernden und anspruchsvolleren Kundenbedürfnissen. Ein Kunde erwartet heute ein nahtloses Zusammenspiel von Reisebüros, Fluggesellschaften, Hotels, Schifffahrtsgesellschaften und anderen Transportunternehmen bei seinen Geschäfts- oder Ferienreisen.

Angesichts der sehr heterogenen IT-Landschaften in den meisten Unternehmen, nämlich einer Mischung aus Jahrzehnte alten Legacy-Systemen und modernsten Internetanwendungen, dem extremen Wettbewerb zwischen den Reiseanbietern und immer neuen regulatorischen Auflagen ergeben sich hohe Anforderungen an die Reaktionsfähigkeit der Informatiksysteme. Sie müssen laufend den Änderungen der Geschäftsmodelle und der Geschäftsprozesse angepasst werden, zum Beispiel bei der Bildung von Allianzen oder der Lancierung neuer Angebote – und das mit möglichst geringen Kosten.

Vor diesem Hintergrund kann das Paradigma der Service-orientierten Architektur eine entscheidende Grundlage für die Modernisierung bestehender Software und den Bau neuer Applikationen werden und damit die Wettbewerbsfähigkeit der Unternehmen wirkungsvoll unterstützen. Siehe auch [Bu07].

EDS hat als führender IT Service Provider für die Transportindustrie vor einigen Jahren begonnen eine entsprechende Strategie für ihre Kunden zu entwickeln und in ein umfassendes Angebot investiert, die „EDS Airline SOA“. Dabei wurde von der Konzernstrategie der „Agile Enterprise Plattform“ ausgegangen und die Erfahrungen mit SOA aus anderen Branchen, z.B. der Finanzindustrie einbezogen. [KN07]

3 Die EDS Airline SOA

3.1 Agile Enterprise Plattform

Die „EDS Agile Enterprise Plattform“ ist eine Netzwerk-basierte Plattform, welche eine flexible Technologie-Basis schafft, die den Unternehmen hilft rasch und kosteneffizient auf Marktanforderungen zu reagieren und damit ihre Wettbewerbsfähigkeit zu steigern. Dabei ist Service-Orientierung eines der wichtigen Entwurfsprinzipien für Agilität, wie in Abbildung 1 dargestellt.

Die Service-Orientierung unterstützt als Integrationsparadigma auch direkt die Vorstellung, die mit dem Begriff „Digital Nervous System“ verbunden ist. Dabei wird die Informationsverarbeitung im Unternehmen mit der Funktionsweise des Nervensystems des Menschen verglichen. Diese Analogie wurde unter anderem von Bill Gates vor etwa 10 Jahren populär gemacht. Genau so wie ein lebender Organismus ununterbrochen sensorische und andere Informationen erfasst, verarbeitet und damit Aktionen auslöst, soll entscheidungsorientierte Informationsverarbeitung im Unternehmen nahtlos integriert funktionieren.

EDS Agile Enterprise Platform

SOA & Event Driven “DNS” for the Airline Industry

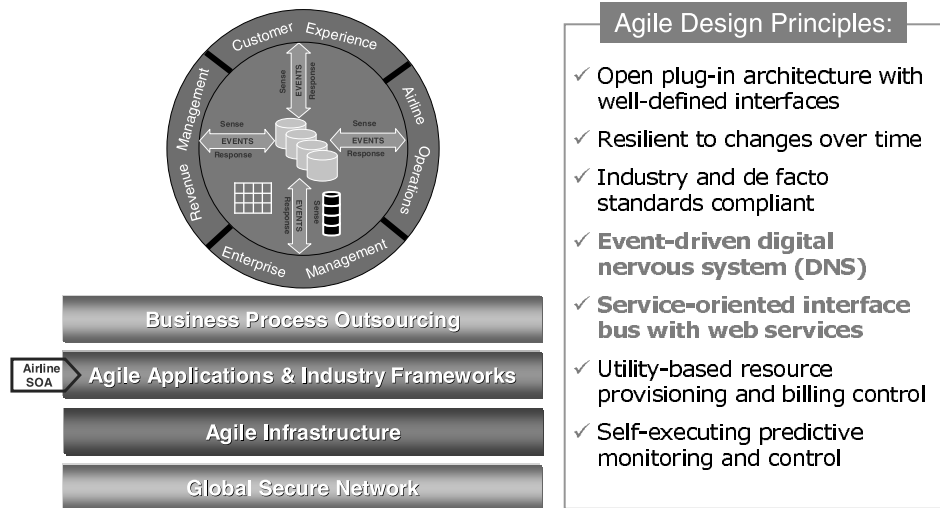


Abbildung 1: EDS Agile Enterprise Plattform

Die EDS SOA Strategie ist Teil des umfassenden Angebots der „Agile Enterprise Plattform“. Besonderes Augenmerk liegt dabei auf den Gesamtkosten (TCO Total Cost of Ownership) des SOA Vorhabens, zusammengefasst bezeichnet als “Designed for Run™”. Gemeint ist die Situation, dass üblicherweise das Augenmerk häufig nur auf der Entwicklungsphase liegt, statt alle Kosten einer voll funktionsfähigen Lösung im voraus zu berücksichtigen. Heutige Architekturen verlangen beispielsweise robuste Testumgebungen, welche den ausfallfreien 24-Stunden Betrieb realistisch nachbilden. Die vielen Elemente einer anspruchsvollen SOA Architektur müssen einwandfrei zusammenarbeiten, sonst ist der unterstützte Geschäftsprozess gefährdet.

Eine der ersten Gelegenheiten, die ganze Bandbreite der Erfahrungen einer vollständigen geschäftskritischen Lösung (Hardware, Netzwerk und Applikationen) zu machen, bot sich EDS in der Luftfahrtindustrie. [WS08]

Warum hilft eine SOA einem Unternehmen bzw. der IT des Unternehmens agiler zu werden? Weil die Interoperabilität und die Wiederverwendung von Standard-Services die Flexibilität erhöht und die Integrationsaufgaben erleichtert. Gleichzeitig werden Investitionen in bewährte, robuste Legacy-Systeme geschützt. Modernisierung dieser Legacy-Systeme ist deshalb ein Schlüssel der Transformation der IT Landschaft, um diese Funktionalität als Service im Rahmen der SOA nutzen zu können.

3.2 **Airline SOA Strategie**

Treibende Kräfte für den Wandel der IT in Richtung SOA sind im Fall von Fluggesellschaften typischerweise Prozesse für Interline Electronic Ticketing, Frequent Flyer Bonus Programme, Allianzen und Zusammenarbeitsabkommen aller Art zwischen Reise-dienstleistern, neue gesetzliche Vorgaben, etwa im Sicherheitsbereich, neue Vertriebskanäle, z.B. Buchung über Mobilgeräte und umfassende, präzise Informations-systeme für den Kunden. Verglichen mit der vollständigen Ablösung aller Systeme durch eine neue Generation in einem Zug, stellt der SOA Ansatz eine interessante Alternative für eine schrittweise Erneuerung derselben Systeme dar.

Wie geht man nun bei der Entwicklung einer SOA Strategie konkret vor?

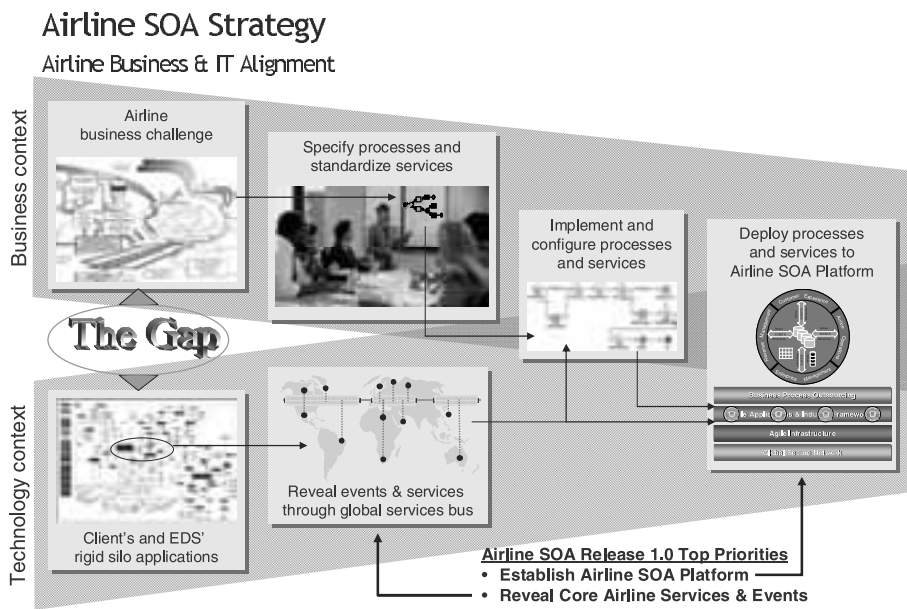


Abbildung 2: Airline SOA Strategie

Grundsätzlich geht man bei der Entwicklung einer SOA Lösung von den kritischen Geschäftsprozessen der Fluggesellschaft einerseits (Abbildung 2 oben), und den bestehenden Applikationen und Daten (Abbildung 2 unten) andererseits aus. Aufgrund einer Analyse der neuen, zukünftigen Anforderungen an die Geschäftsprozesse wird ein Plan entwickelt, die Lücken schrittweise zu schliessen (im Bild „The Gap“, Abstimmung IT und Geschäft [Fe07]). Normalerweise besteht die Kernaufgabe darin, die Integrationsschicht zu definieren (Enterprise Service Bus) und die bestehenden Applikationen, soweit sie zukunftsfähig sind, als standardisierte Services zu exponieren.

Welche Anforderungen und Herausforderungen nun im konkreten Falle der Lufthansa zu bewältigen waren, wird im Folgenden dargestellt.

4 Ausgangslage: Lufthansas Common IT Platform

Der Auslöser für die Umsetzung des SOA Projektes ist auf Lufthansas Entscheid zurückzuführen, auf ein neues von der Firma Amadeus entwickeltes Passenger Service System (PSS) zu wechseln. In einem solchen System werden nicht nur sämtliche Flugbuchungen und deren Änderungen verwaltet (Reservierungssystem), sondern auch die Sitzplatzverfügbarkeiten und die Gesamtauslastung eines Fluges (Inventarsystem), sowie die Passagier- und Gepäckabfertigung (Check-In System). Wie man daraus leicht erkennen kann fliessen eine grosse Anzahl an Inputparametern, wie z.B. Passagier- und Frequent Traveller Daten, Preise, Buchungsstände und Sitzplatzauslastungen, aber auch Informationen zu Flugplanänderungen und Anschlussflügen, um nur einige zu nennen, in einem solchen System zusammen. Diese Daten werden durch unterschiedliche externe Anwendungen gesteuert und müssen mit diesen kontinuierlich ausgetauscht werden. In einer ersten Phase sind etwa 30 Lufthansa Anwendungen betroffen, welche Daten mit dem Amadeus PSS System (und umgekehrt) austauschen. Viele dieser Anwendungen benutzen unterschiedliche Datenformate und Transportmethoden, was schnell dazu führen kann, dass viele Konvertierungsschritte von Nöten sind, um wieder miteinander kommunizieren zu können.

Aus dieser Situation heraus hat Lufthansa, als eines der ersten Star Allianz Mitglieder den bedeutenden Strategieentscheid gefällt, eine Middleware Plattform zwischen der Amadeus „Common IT Platform“ (CITP, siehe Abbildung 3) und den Lufthansa Anwendungen aufzubauen. Dabei sollte diese Plattform bereits im ersten Schritt so ausgelegt werden, dass sie später auch den anderen Star Allianz Mitgliedern, ohne grossen Zusatzaufwand, Zugang zu den gemeinsam genutzten Systemen gewähren kann. Damit war der Grundstein für die Middleware gelegt und das Lufthansa SOA (LH SOA) Projekt geboren.

Aber was sind denn nun konkret die grossen Vorteile einer solchen Plattform?

Common IT Platform will enable consistent, seamless service across Star Alliance

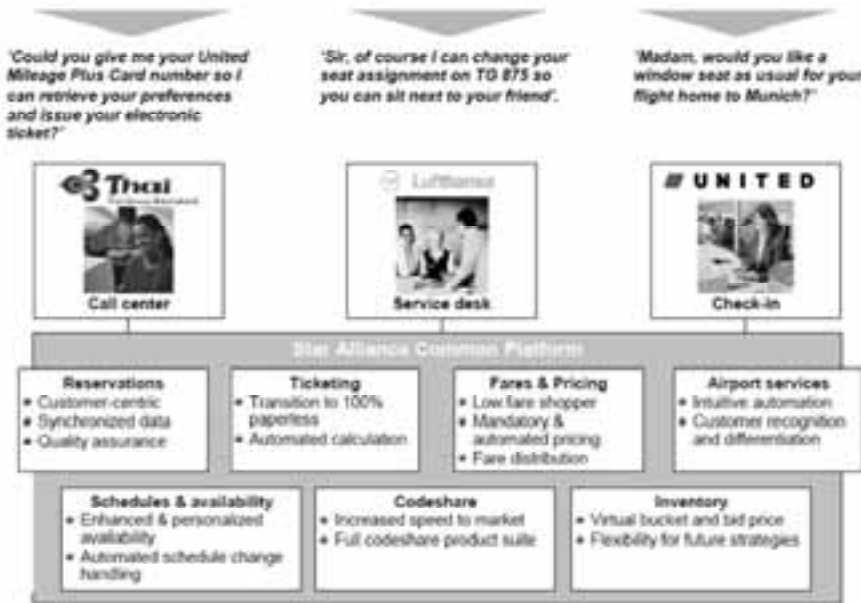


Abbildung 3: STAR Alliance Common IT Platform
(mit freundlicher Genehmigung von Lufthansa)

Mit der neuen gemeinsamen Plattform sind alle daran teilnehmenden Star Carrier (Fluggesellschaften) in der Lage harmonisierte und integrierte Services anzubieten. Dies führt dazu, dass die Carrier die Prozesse ihrer Hauptanwendungen gemeinsam nutzen können, gleichzeitig jedoch weiterhin ihre separaten Prozesse für spezifische Anwendungen beibehalten können. So gesehen verbessert CITP die Business Situation der gesamten Star Allianz wesentlich, was zu einem Wettbewerbsvorteil gegenüber den Konkurrenten führt.

Die zahlreichen Vorteile liegen auf der Hand:

- *Kostenreduktion:* Vor allem die IT Kosten in den wichtigen Bereichen wie Betrieb, Entwicklung und Wartung werden dadurch signifikant reduziert. Durch den Synergie-Effekt der gemeinsamen Nutzung kommt es zu weiteren Kosteneinsparungen für alle Beteiligten. [Mc06]
- *Vereinheitlichung der Services:* Durch die gemeinsame Nutzung der CITP ist es möglich Prozesse und Produkte der Star Carrier zu vereinheitlichen und so dem Kunden ein harmonisiertes Produkt anzubieten. Zudem werden Geschäftsprozesse wie beispielsweise das Durchchecken von Passagieren und Gepäck über mehrere Flugsegmente und Fluggesellschaften erleichtert.

- *Schnellere Markteinführung (time to market)*: Dank der CITP ist es möglich schneller neue Funktionalitäten und Produkte einzuführen. Diese können dann von allen teilnehmenden Carriern gemeinsam benutzt werden.

Basierend auf moderner Technologie und einer State-of-the-Art Implementierung unterstützt die CITP, dank der neu zur Verfügung stehenden Flexibilität bei reduzierten Kosten, die zukunftssträchtige IT Strategie der Lufthansa.

5 Zielsetzung des LH SOA Projektes und seine nicht-funktionalen Anforderungen

Ausgehend von der bestehenden EDS Airline SOA Plattform hat EDS für Lufthansa eine Middleware Plattform entwickelt, über welche die Anbindung verschiedener Lufthansa Backend-Systeme ermöglicht wird und welche somit zum zentralen Bestandteil der CITP wurde.

Um die Lufthansa Applikationen mit Altea, der Amadeus PSS Lösung, zu verbinden, wurden drei verschiedene Kommunikationskanäle (synchron, asynchron und Batch) zur Verfügung gestellt, deren Implementierung im nächsten Kapitel näher beschrieben wird.

Im Folgenden werden kurz einige der Wichtigsten nicht-funktionalen Anforderungen an die SOA Plattform aufgeführt:

Synchrone Kommunikation:

- Volumen von 800 Millionen Transaktionen pro Monat
- Spitzenwerte von 360 zustandslosen (stateless) Serviceaufrufen pro Sekunde
- Skalierbarkeit auf Spitzenwerte von über 720 zustandslosen Serviceaufrufen pro Sekunde ohne Architekturanpassungen

Asynchrone Kommunikation:

- Volumen von 230 Millionen Meldungen pro Monat
- Spitzenwerte von 300 Meldungen pro Sekunde
- Skalierbarkeit auf Spitzenwerte von über 540 Meldungen pro Sekunde ohne Architekturanpassungen

Weitere Kennzahlen:

- Middleware Verfügbarkeit von 99.95% pro Monat
- 100%-ige Redundanz (keine „single points of failure“)
- Unterstützung von 10,000 gleichzeitigen Nutzern
- Latenzzeit von max. 30ms (hin und zurück) für 90% der typischen synchronen Serviceaufrufe

6 Technischer Lösungsansatz

Die Lufthansa SOA Plattform basiert auf dem EDS Airline SOA Design. Sie ist eine Service-orientierte Middleware Plattform, welche Systemintegration und Kommunikation über synchrone und asynchrone sowie Batchverarbeitung unterstützt. Ebenso umfasst sie Event gesteuerte Prozesse, Benachrichtigungsdienste und Scheduling Services für Files. Über eine grosse Anzahl von Schnittstellen und Adaptoren, welche über einen Orchestrierungsmechanismus (Tibco BW Orchestration Engine) gesteuert werden, werden verschiedene Lufthansa Applikationen mit dem Amadeus Altea System verbunden. Alle Web services wurden mit WSDL beschrieben. BPEL wurde nicht verwendet.

Die Architektur (Abbildung 4) und ihre Komponenten erlauben flexible Umsetzungsstrategien, um Performance und Hardwarenutzung zu optimieren, ermöglichen die logische und physische Separierung der Geschäftslogik vom ‚Kern‘ der Middleware, sowie verteilte Zuständigkeiten und Management.

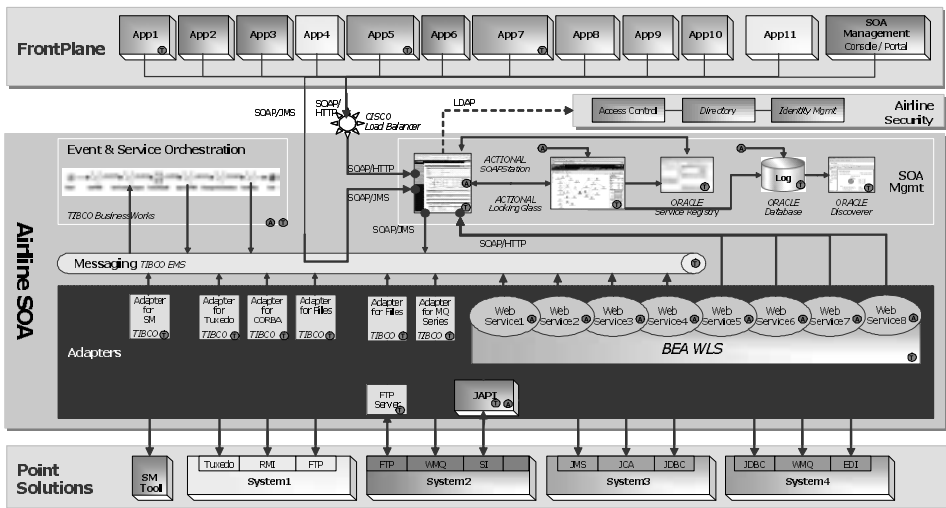


Abbildung 4: Überblick LH SOA Architektur

LH SOA Komponenten

Wie aus der obigen Darstellung ersichtlich ist, besteht die Lufthansa SOA aus einer Vielzahl von Komponenten von Drittanbietern sowie kundenspezifischen Entwicklungen, welche über die Plattform nahtlos integriert werden. Die Hauptkomponenten sind:

Actional SOAPstation und LookingGlass

SOAPstation (SST) stellt eine sichere Verteilung von Web Services zu verschiedenen Anwendungen her. Gleichzeitig dient sie auch als Proxy zwischen den Applikationen, welche Web Services zur Verfügung stellen und den Systemen, die darauf aufbauen. Die wichtigsten Vorteile sind hierbei die Bereitstellung von Services über verschiedene Protokolle, die Zugangskontrolle über eine Kombination von Credentials, Business Roles

und Rules, sowie die Versions- und Änderungskontrolle. SST liefert Logs und Reports über die Serviceaktivität sowohl auf Betriebs- als auch auf Geschäftsebene. Dies ermöglicht ein durchgängiges Service Level Management und Überwachung der verschiedenen Aktivitäten.

GhostAgents run continuously on each application server on the network that hosts web services detecting each incoming and outgoing message. The agents collect and evaluate run-time statistics and implement policies. The GhostAgents do not affect the runtime message flow and clients and servers are unaware that they exist.

GhostAgents laufen auf jedem Applikationsserver im Netz, welcher Web services hostet und detektieren alle Nachrichten. Diese Agenten sammeln und evaluieren Laufzeit-Statistiken und implementieren Regeln. Sie sind für die clients und server unsichtbar

Der LookingGlass (LG) Server ist die zentrale Engine, welche Nachrichtenflüsse, die von den GhostAgents kommen korreliert und Policies an diese Agenten verteilt. LG ermöglicht einen horizontalen Einblick in den Nachrichtenfluss innerhalb der Middleware und erlaubt so einem Administrator, Netzwerk und Daten zu analysieren und systemweite Policies zu erstellen. LG ermittelt darüber hinaus Abhängigkeiten von tatsächlichen Web Service Laufzeiten und generiert Performance Statistiken. Es ermöglicht auch gezielte Ursachenforschung bei Problemen oder Systemausfällen.

TIBCO Enterprise Service Bus

Die zentralen Nachrichtenflüsse auf Unternehmensebene wurden mittels TIBCO Enterprise Message Server (EMS), Business Works (BW) und Adaptern implementiert.

EMS liefert JMS 1.1 konforme fehlertolerante und persistente Nachrichtendienste, und – wo benötigt – Sicherheitsdienste mit einer Authentifizierung gegen ein LDAP im Unternehmens. Lastbegrenzungsmechanismen werden über Nachrichtenverfall und –steuerung ermöglicht.

BW stellt Orchestrierung, Transformation und Protokollwechsel-Funktionalitäten zur Verfügung, um die verschiedener Client und Backend-Applikationen und Prozesse zu integrieren und den Datenflusses zwischen verschiedenen Applikationen und Systemen sicherzustellen.

BW Adapter sorgen auf oberster Ebene für die Integration und Interoperabilität mit bestehenden Altapplikationen wie Unisys Mainframe oder MQ Series.

TIBCO Hawk und Administrator

TIBCO Hawk ist ein Werkzeug für das Monitoring und Management verteilter Applikationen und Betriebssysteme innerhalb der LH SOA, unter Verwendung der TIBCO Messaging Infrastruktur für die Kommunikation. Hawk ist ein ereignisbasiertes System, aufgebaut auf dem Konzept eines verteilten, autonomen, intelligenten Agenten, welcher auf jeder verwalteten Maschine im Netzwerk läuft. Es wurden über 200 LH SOA-spezifische Hawk Regelwerke implementiert, um einen nahtlosen Betrieb der Plattform

sicherzustellen und die strengen Hochverfügbarkeitsanforderungen zu erfüllen. Hawk meldet alle Ereignisse und Statusinformationen an den zentralen TIBCO Administrator, welcher über eine grafische Benutzeroberfläche das Management der LH SOA Komponenten ermöglicht.

BEA WebLogic

BEA WebLogic Server ist ein Applikationsserver für den Java 2 Enterprise Edition (J2EE) Standard welcher die Informationen und Funktionalitäten der Amadeus und LH Backend-Systeme für den Service Bus als Web Service exponiert.

GlobalScape

GlobalScape Enhanced File Transfer Server (EFT) stellt sichere und nicht gesicherte FTP Dienste zwischen Amadeus und LH zur Verfügung.

Oracle RAC Database (DB) und Application Server (AS)

In der LH SOA wird die ORACLE RAC 10g Datenbank eingesetzt, um Nutzungsstatistiken und Audit-Informationen zu protokollieren und Berichte zur Verfügung zu stellen. Die in der Datenbank gespeicherten Informationen umfassen Statistiken und Auditinträge (logs), welche von den verschiedenen Komponenten produziert werden, z.B. von SST, LG und Hawk, sowie Metadatenverzeichnisse der verschiedenen SOA Komponenten, wie z.B. von Actional, TIBCO, ORACLE Portal, sowie dem Online-Reporting.

Andere ORACLE Produkte werden für das Management Portal, das Service Verzeichnis, LDAP und die Bereitstellung von Business Intelligence Funktionen in der LH SOA genutzt.

LH SOA Nachrichtenflüsse

Die LH SOA verwendet drei Hauptkanäle für den Nachrichtenfluss, die von den verschiedenen Arten von Clients und Backend-Applikationen genutzt werden.

Synchroner Transport-Kanal

Der Synchrone Transport-Kanal (synchrone Kommunikation) liefert HTTP/S und JMS Anbindungen für alle Web Service Aufrufe – hierbei dient die SST als zentraler Web Service Proxy. Die eigentlichen Web Service Lieferanten befinden sich hinter der SST und werden üblicherweise mit BEA WebLogic oder in TIBCO Business Works implementiert. Der Einsatz der SST entkoppelt so die Clients von den Backend-Systemen.

Clients können HTTP oder JMS-basierte Transport-Protokolle nutzen, jeweils mit oder ohne Secure Service Layer (SSL), wie in der folgenden Grafik dargestellt. Die Verfahrensregeln (Policies) für Web Services werden in SST konfiguriert, z.B. für Lastverteilung und Sicherheit (Security Policies) etc.

Die Abbildung 5 gibt einen Überblick über den synchronen Nachrichtenfluss: Anfragen treffen über HTTP oder EMS ein, werden von der SST verarbeitet, welche wiederum die benötigten BEA oder BW Web Services aufruft, die dann die zugehörigen Backend JAPI Objekte ansprechen. Die Backend-Antworten werden schliesslich der anfragenden Applikation via BEA und SST, entweder über HTTP/S oder eine temporäre Queue via EMS zurückgemeldet.

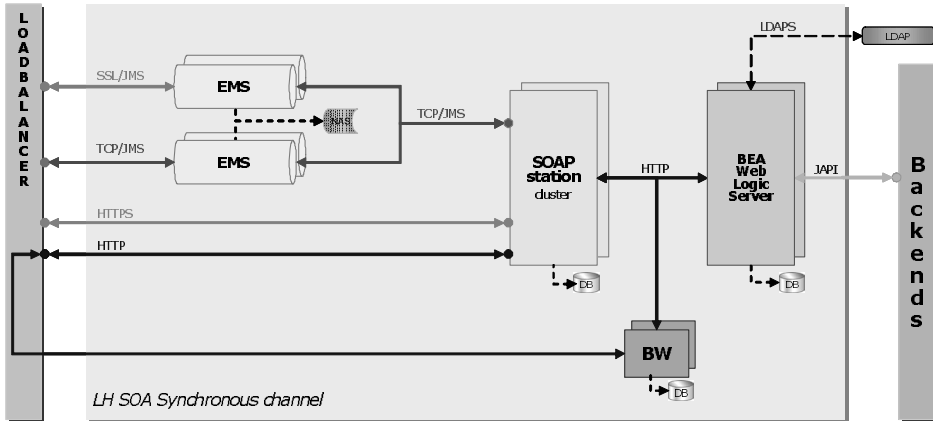


Abbildung 5: LH SOA Implementierung synchrone Kommunikation

Asynchroner Transport-Kanal

Die LH SOA unterstützt auch asynchrone Kommunikationsdienste. Eine Interaktion wird „asynchron“ genannt, wenn die den Service aufrufende Applikation nicht auf die Verfügbarkeit des Resultates der Operation wartet, oder wenn überhaupt kein Resultat erwartet wird. Der asynchrone Transport-Kanal der LH SOA unterstützt offene Standards, wie z.B. JMS, Punkt-zu-Punkt oder Publish/Subscribe Verfahren, sowie verschiedene Sicherheitsdienste. Ebenso ist er skalierbar und redundant konfiguriert, d.h. ohne „single points of failure“

Eine externe Applikation sendet eine Nachricht zu einer spezifischen Warteschlange (Queue) im Nachrichten System. Der Name dieser Warteschlange bezeichnet dabei jeweils die auszuführende Geschäfts- oder Integrationslogik. Die SOA Middleware erkennt eingehende Nachrichten in einer spezifischen Warteschlange und löst einen BW Integrationsprozess und die jeweilige Service Funktionalität aus.

Sofern der Service ein Ergebnis an die aufrufende Applikation oder an den Client zurücksenden soll, so erstellt der BW Prozess eine entsprechende Ergebnismeldung und sendet diese Nachricht an eine Ausgangswarteschlange. Der Client, der die Warteschlange „abhört“, ruft die Meldung ab, sobald sie in der Warteschlange erscheint. Der asynchrone Nachrichtenfluss ist in Abbildung 6 dargestellt.

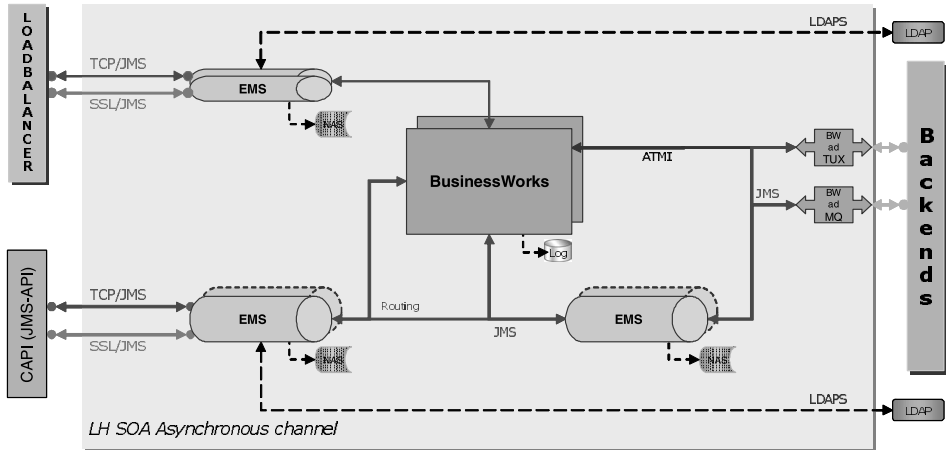


Abbildung 6: Implementierung asynchrone Kommunikation

Batch Transport-Kanal

Batch Prozesse werden über eine sichere FTP Umgebung innerhalb der LH SOA verarbeitet. Ein Client sorgt dabei für das Senden (Send) und Empfangen (Receive) von Dateien von externen Systemen. Dabei findet das Herauf- oder Herunterladen nie direkt mit dem externen System statt, sondern immer über die LH SOA. Die Middleware kümmert sich dabei um die regelbasierte Weiterleitung der Daten an die anderen Systemen. Dateien werden dabei von den externen Systemen auf dem Middleware FTP Server geladen. Die Middleware leitet sie dann zum Ziel-Client weiter oder bietet sie diesem zum Herunterladen an.

LH SOA Management

Um einen nahtlosen, hochverfügbaren Betrieb sicher zu stellen, ist ein effizientes Monitoring aller zentralen SOA Komponenten der LH SOA Plattform ein entscheidender Faktor. Dies wird durch LH SOA Management Werkzeuge auf verschiedenen Ebenen ermöglicht:

Horizontal: Nachrichtenflüsse werden durch Actional SOAPstation und Ghost Agents gesteuert

Vertikal: Das zentrales Management der einzelnen Komponenten wurde mit TIBCO Hawk Agents umgesetzt.

Das LH SOA Managementkonzept wird in der folgenden Abbildung 7 dargestellt:

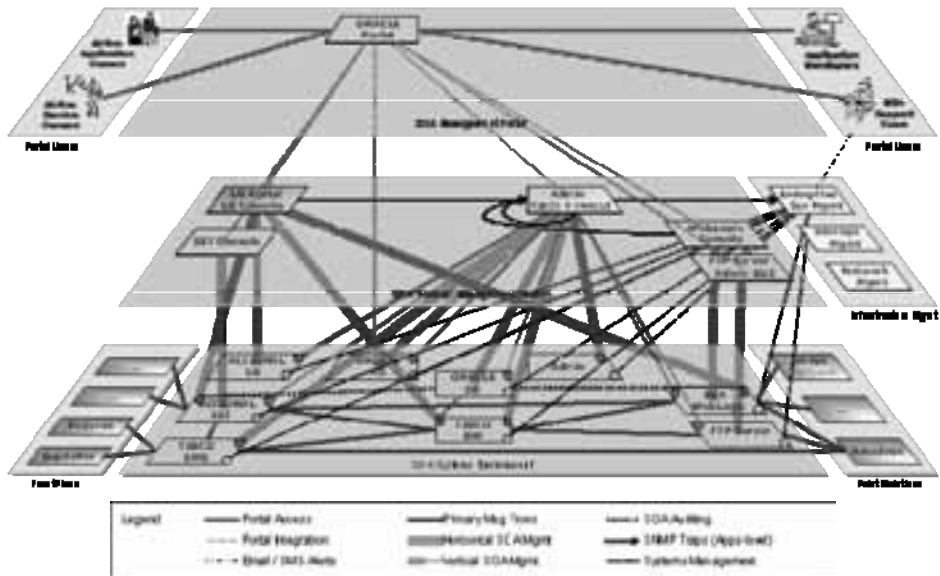


Abbildung 7: Überblick LH SOA Management

1. Die „SOA Systemumgebung“ (unterste Ebene) umfasst diejenigen Hardware und Software Komponenten, welche die SOA Runtime Plattform bilden und für den Nachrichtenfluss sorgen.
2. Das „SOA Platform Management Toolset“ (mittlere Ebene) liefert die Funktionalitäten für die horizontale SOA Nachrichtensteuerung und das vertikale SOA Komponenten Management. Darüber hinaus gibt es auf der Infrastrukturseite ebenfalls Werkzeuge für die Steuerung und Überwachung der Infrastruktur Services.
3. Das „SOA Management Portal“ (oberste Ebene) bildet das „Cockpit“ der SOA mit einer einheitlichen Benutzeroberfläche für das Monitoring und die Administration der gesamten SOA Plattform. Primäre Anwendergruppen hierfür sind die Verantwortlichen für die jeweiligen Dienste oder Applikationen, sowie Entwicklungs- und Betriebsteams.

7 Erfahrungen aus der Projekt Implementierung

Im folgenden Kapitel werden die wichtigsten Herausforderungen und Erfahrungen des LH SOA Implementierungsprojektes beschrieben.

7.1 Herausforderungen

Was das Design der Middleware angeht, so wurde auf dem bereits existierenden EDS Airline SOA Release aufgebaut, welcher jeweils mit den zusätzlichen Lufthansa Anforderungen ergänzt wurde. Trotzdem blieb ein hoher Level an technologischer Komplexität erhalten, welcher sich nicht zuletzt auch im Zusammenspiel der eingesetzten Produkte zeigte.

So ist es nicht erstaunlich, dass sich das *Provider Management* als eine der zeitaufwändigsten Aufgaben im Projekt herausstellte. Neben dem regulären Provider Tracking, welches man in jedem Projekt mit Drittanbietern wiederfindet, gab es im LH SOA Projekt vor allem technologische Herausforderungen, welche sich primär an den Schnittstellen zwischen den verschiedenen eingesetzten Produkte bemerkbar machten. Auch, dass viele der Drittanbieter Konkurrenten waren, kam erschwerend hinzu. So waren im Laufe des Projektes mehrere Task Forces von Nöten, in welchen die beteiligten Anbieter an einen Tisch zusammengebracht werden mussten, um eine konstruktive Lösung der aufgetauchten Probleme gemeinsam herbeizuführen.

Als zweite wichtige Erfahrung in diesem Projekt möchten wir das Thema *Testen* herausgreifen. So wurden die Testaufwände in der ursprünglichen Planung dieses Projektes stark unterschätzt, wobei noch hinzuzufügen ist, dass EDS im Auftrag der Lufthansa auch die Abnahmetests durchgeführt hat. Die hohen Aufwände sind auf verschiedene Faktoren zurückzuführen: Zum einen wurde der Aufwand für die Erstellung der Testfälle (es gab ca. 500 davon) und Abnahmekriterien, sowie die generelle Dokumentation der Testläufe unterschätzt. Zum anderen war vor allem der Aufwand für die Durchführung der Hochverfügbarkeitstests (Failover und Redundanz) wesentlich höher als erwartet. Jeder einzelne dieser Tests brachte grosse Planungs- und Koordinationsaufwände mit sich, beginnend mit der Verfügbarkeit aller involvierten Parteien über die Testfenster der Testplattformen bis hin zur eigentlichen Durchführung der komplexen Testfälle, welche in einigen Fällen für einen einzelnen Test alleine bereits 8 Stunden beanspruchte. Anschliessend folgten die Auswertung der Testergebnisse und allenfalls Suche nach Fehlerursachen sowie deren Behebung. Dieser ganze Zyklus musste je nach Testfall zum Teil mehrmals wiederholt werden.

Zusammenfassend lässt sich sagen, dass das Testen (inkl. gewisser Aufwände für das Anforderungsmanagement, das Erstellen der Abnahmekriterien und die Fehlerkorrektur) etwa zwischen $\frac{1}{4}$ und $\frac{1}{3}$ des Gesamtaufwandes des Implementierungsprojektes ausgemacht hat.

7.2 Erfahrungen & kritische Erfolgsfaktoren

Abschliessend möchten wir einige der wichtigsten Erfolgsfaktoren zusammenfassen:

- Sorgfältige Reviews der Projektaufwände während der Planungsphase mit Fokus auf die technologische Komplexität und die zu erwartenden Testaufwände

- Architekturdesign auf dem richtigen Granularitätslevel, welches vor allem das Potential für spätere Wieder- und Weiterbenutzung berücksichtigt
- Setup einer Projektmanagement Organisation, welche in der Lage ist gleichermaßen effizient die global verteilte Stakeholder, Drittanbieter sowie On- und Off-Shore Teams zu führen
- Sicherstellung bzw. rascher Aufbau einer adäquaten technologischen Wissensbasis mit vertieftem Anwendungs-Know-how bei allen Beteiligten
- Proaktive Teilnahme an Industrie Foren
- Motivation des Projektteams über einen langen Zeitraum und insbesondere in schwierigen Projektphasen
- Zielorientierte Führungskräfte, die sich der Komplexität und technologischen Herausforderungen eines solchen Projektes bewusst sind
- Definition und Umsetzung geeigneter Führungsgremien (Governance-Boards)
- Zentrales Produktmanagement der SOA Plattform inklusive aller damit verbundenen Themenbereiche

8 Zusammenfassung und Ausblick

Das vorliegende Beispiel illustriert die erfolgreiche Realisierung und das Nutzenpotential eines SOA Ansatzes. Eine Service-orientierte Architektur hilft die Forderung nach Flexibilität und gleichzeitiger Kosteneffizienz unter einen Hut zu bringen, indem gemeinsam genutzte Services durch Integrations- und Kommunikationsmechanismen verbunden werden. Auch in diesem Projekt zeigt sich der Trend, dass der Nutzen von SOA direkt von der Leistungsfähigkeit der Integrationsebene, d.h. der Middleware, bestimmt wird und erfolgreiche SOA Ansätze stark Nachrichtenfluss-zentriert sind.

In diesem Sinne hilft eine „agile IT“ einem Unternehmen „agil“ zu werden, da sowohl Kostenoptimierung als auch eine verbesserte Reaktionsfähigkeit bei der Weiterentwicklung von Geschäftsprozessen gleichzeitig angestrebt werden können.

Das SOA Projekt mit der Lufthansa zeigt den Weg in die Zukunft. So haben bereits weitere Fluggesellschaften Interesse angemeldet, für welche EDS derzeit eine modifizierte Version der EDS Airline SOA implementiert.

9 Danksagung

Die Autoren danken dem Architekturteam unter der Leitung von Peter Schmid für die Unterstützung der Ausarbeitung dieses Beitrags.

10 Literaturverzeichnis

- [Bu07] Burger, L., SOA IN THE TRANSPORTATION INDUSTRY, EDS Viewpoint Paper, 2007.
- [Fe07] Are You Really Ready for SOA? CIO, http://www.cio.com/article/121850/Are_You_Really_Ready_for_SOA_
- [KN07] Kubli R., Nadhan E.G., BANKING ON AN SOA EVOLUTION, EDS Viewpoint Paper, 2007.
- [Mc06] Lufthansa IT Chief Looks To Cut Costs With SOA Software, InformationWeek, <http://www.informationweek.com/story/showArticle.jhtml?articleID=195300016>
- [Wi06] Wikipedia, http://de.wikipedia.org/wiki/Serviceorientierte_Architektur
- [WS08] Interview, itp, SOA Roundtable, <http://www.itpreport.com/default.asp?Mode=Show&A=1314&R=G>

GI-Edition Lecture Notes in Informatics

- P-1 Gregor Engels, Andreas Oberweis, Albert Zündorf (Hrsg.): Modellierung 2001.
- P-2 Mikhail Godlevsky, Heinrich C. Mayr (Hrsg.): Information Systems Technology and its Applications, ISTA'2001.
- P-3 Ana M. Moreno, Reind P. van de Riet (Hrsg.): Applications of Natural Language to Information Systems, NLDB'2001.
- P-4 H. Wörn, J. Mühling, C. Vahl, H.-P. Meinzer (Hrsg.): Rechner- und sensorgestützte Chirurgie; Workshop des SFB 414.
- P-5 Andy Schürr (Hg.): OMER – Object-Oriented Modeling of Embedded Real-Time Systems.
- P-6 Hans-Jürgen Appelrath, Rolf Beyer, Uwe Marquardt, Heinrich C. Mayr, Claudia Steinberger (Hrsg.): Unternehmen Hochschule, UH'2001.
- P-7 Andy Evans, Robert France, Ana Moreira, Bernhard Rumpe (Hrsg.): Practical UML-Based Rigorous Development Methods – Countering or Integrating the extremists, pUML'2001.
- P-8 Reinhard Keil-Slawik, Johannes Magenheimer (Hrsg.): Informatikunterricht und Medienbildung, INFOS'2001.
- P-9 Jan von Knop, Wilhelm Haverkamp (Hrsg.): Innovative Anwendungen in Kommunikationsnetzen, 15. DFN Arbeitstagung.
- P-10 Mirjam Minor, Steffen Staab (Hrsg.): 1st German Workshop on Experience Management: Sharing Experiences about the Sharing Experience.
- P-11 Michael Weber, Frank Kargl (Hrsg.): Mobile Ad-Hoc Netzwerke, WMAN 2002.
- P-12 Martin Glinz, Günther Müller-Luschnat (Hrsg.): Modellierung 2002.
- P-13 Jan von Knop, Peter Schirmbacher and Viljan Mahni_ (Hrsg.): The Changing Universities – The Role of Technology.
- P-14 Robert Tolksdorf, Rainer Eckstein (Hrsg.): XML-Technologien für das Semantic Web – XSW 2002.
- P-15 Hans-Bernd Bludau, Andreas Koop (Hrsg.): Mobile Computing in Medicine.
- P-16 J. Felix Hampe, Gerhard Schwabe (Hrsg.): Mobile and Collaborative Business 2002.
- P-17 Jan von Knop, Wilhelm Haverkamp (Hrsg.): Zukunft der Netze –Die Verletzbarkeit meistern, 16. DFN Arbeitstagung.
- P-18 Elmar J. Sinz, Markus Plaha (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2002.
- P-19 Sigrid Schubert, Bernd Reusch, Norbert Jesse (Hrsg.): Informatik bewegt – Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 30.Sept.-3.Okt. 2002 in Dortmund.
- P-20 Sigrid Schubert, Bernd Reusch, Norbert Jesse (Hrsg.): Informatik bewegt – Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 30.Sept.-3.Okt. 2002 in Dortmund (Ergänzungsband).
- P-21 Jörg Desel, Mathias Weske (Hrsg.): Promise 2002: Prozessorientierte Methoden und Werkzeuge für die Entwicklung von Informationssystemen.
- P-22 Sigrid Schubert, Johannes Magenheimer, Peter Hubwieser, Torsten Brinda (Hrsg.): Forschungsbeiträge zur “Didaktik der Informatik” – Theorie, Praxis, Evaluation.
- P-23 Thorsten Spitta, Jens Borchers, Harry M. Sneed (Hrsg.): Software Management 2002 – Fortschritt durch Beständigkeit
- P-24 Rainer Eckstein, Robert Tolksdorf (Hrsg.): XMIDX 2003 – XML-Technologien für Middleware – Middleware für XML-Anwendungen
- P-25 Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Commerce – Anwendungen und Perspektiven – 3. Workshop Mobile Commerce, Universität Augsburg, 04.02.2003
- P-26 Gerhard Weikum, Harald Schöning, Erhard Rahm (Hrsg.): BTW 2003: Datenbanksysteme für Business, Technologie und Web
- P-27 Michael Kroll, Hans-Gerd Lipinski, Kay Melzer (Hrsg.): Mobiles Computing in der Medizin
- P-28 Ulrich Reimer, Andreas Abecker, Steffen Staab, Gerd Stumme (Hrsg.): WM 2003: Professionelles Wissensmanagement – Erfahrungen und Visionen
- P-29 Antje Düsterhöft, Bernhard Thalheim (Eds.): NLDB'2003: Natural Language Processing and Information Systems
- P-30 Mikhail Godlevsky, Stephen Liddle, Heinrich C. Mayr (Eds.): Information Systems Technology and its Applications
- P-31 Arslan Brömmel, Christoph Busch (Eds.): BIOSIG 2003: Biometric and Electronic Signatures

- P-32 Peter Hubwieser (Hrsg.): Informatische Fachkonzepte im Unterricht – INFOS 2003
- P-33 Andreas Geyer-Schulz, Alfred Taudes (Hrsg.): Informationswirtschaft: Ein Sektor mit Zukunft
- P-34 Klaus Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenber, Wolfgang Wahlster (Hrsg.): Informatik 2003 – Innovative Informatikanwendungen (Band 1)
- P-35 Klaus Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenber, Wolfgang Wahlster (Hrsg.): Informatik 2003 – Innovative Informatikanwendungen (Band 2)
- P-36 Rüdiger Grimm, Hubert B. Keller, Kai Rannenber (Hrsg.): Informatik 2003 – Mit Sicherheit Informatik
- P-37 Arndt Bode, Jörg Desel, Sabine Rathmayer, Martin Wessner (Hrsg.): DeLFI 2003: e-Learning Fachtagung Informatik
- P-38 E.J. Sinz, M. Plaha, P. Neckel (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2003
- P-39 Jens Nedon, Sandra Frings, Oliver Göbel (Hrsg.): IT-Incident Management & IT-Forensics – IMF 2003
- P-40 Michael Rebstock (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2004
- P-41 Uwe Brinkschulte, Jürgen Becker, Dietmar Fey, Karl-Erwin Großpietsch, Christian Hochberger, Erik Maehle, Thomas Runkler (Edts.): ARCS 2004 – Organic and Pervasive Computing
- P-42 Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Economy – Transaktionen und Prozesse, Anwendungen und Dienste
- P-43 Birgitta König-Ries, Michael Klein, Philipp Obreiter (Hrsg.): Persistence, Scalability, Transactions – Database Mechanisms for Mobile Applications
- P-44 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): Security, E-Learning, E-Services
- P-45 Bernhard Rumpe, Wolfgang Hesse (Hrsg.): Modellierung 2004
- P-46 Ulrich Flegel, Michael Meier (Hrsg.): Detection of Intrusions of Malware & Vulnerability Assessment
- P-47 Alexander Prosser, Robert Krimmer (Hrsg.): Electronic Voting in Europe – Technology, Law, Politics and Society
- P-48 Anatoly Doroshenko, Terry Halpin, Stephen W. Liddle, Heinrich C. Mayr (Hrsg.): Information Systems Technology and its Applications
- P-49 G. Schiefer, P. Wagner, M. Morgenstern, U. Rickert (Hrsg.): Integration und Datensicherheit – Anforderungen, Konflikte und Perspektiven
- P-50 Peter Dadam, Manfred Reichert (Hrsg.): INFORMATIK 2004 – Informatik verbindet (Band 1) Beiträge der 34. Jahrestagung der Gesellschaft für Informatik e.V. (GI), 20.-24. September 2004 in Ulm
- P-51 Peter Dadam, Manfred Reichert (Hrsg.): INFORMATIK 2004 – Informatik verbindet (Band 2) Beiträge der 34. Jahrestagung der Gesellschaft für Informatik e.V. (GI), 20.-24. September 2004 in Ulm
- P-52 Gregor Engels, Silke Seehusen (Hrsg.): DELFI 2004 – Tagungsband der 2. e-Learning Fachtagung Informatik
- P-53 Robert Giegerich, Jens Stoye (Hrsg.): German Conference on Bioinformatics – GCB 2004
- P-54 Jens Borchers, Ralf Kneuper (Hrsg.): Softwaremanagement 2004 – Outsourcing und Integration
- P-55 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): E-Science und Grid Ad-hoc-Netze Medienintegration
- P-56 Fernand Feltz, Andreas Oberweis, Benoit Otjacques (Hrsg.): EMISA 2004 – Informationssysteme im E-Business und E-Government
- P-57 Klaus Turowski (Hrsg.): Architekturen, Komponenten, Anwendungen
- P-58 Sami Beydeda, Volker Gruhn, Johannes Mayer, Ralf Reussner, Franz Schweiggert (Hrsg.): Testing of Component-Based Systems and Software Quality
- P-59 J. Felix Hampe, Franz Lehner, Key Pousttchi, Kai Ranneberg, Klaus Turowski (Hrsg.): Mobile Business – Processes, Platforms, Payments
- P-60 Steffen Friedrich (Hrsg.): Unterrichtskonzepte für informatische Bildung
- P-61 Paul Müller, Reinhard Gotzhein, Jens B. Schmitt (Hrsg.): Kommunikation in verteilten Systemen
- P-62 Federrath, Hannes (Hrsg.): „Sicherheit 2005“ – Sicherheit – Schutz und Zuverlässigkeit
- P-63 Roland Kaschek, Heinrich C. Mayr, Stephen Liddle (Hrsg.): Information Systems – Technology and its Applications

- P-64 Peter Liggesmeyer, Klaus Pohl, Michael Goedicke (Hrsg.): Software Engineering 2005
- P-65 Gottfried Vossen, Frank Leymann, Peter Lockemann, Wolffried Stucky (Hrsg.): Datenbanksysteme in Business, Technologie und Web
- P-66 Jörg M. Haake, Ulrike Lucke, Djamshid Tavangarian (Hrsg.): DeLFI 2005: 3. deutsche e-Learning Fachtagung Informatik
- P-67 Armin B. Cremers, Rainer Manthey, Peter Martini, Volker Steinhage (Hrsg.): INFORMATIK 2005 – Informatik LIVE (Band 1)
- P-68 Armin B. Cremers, Rainer Manthey, Peter Martini, Volker Steinhage (Hrsg.): INFORMATIK 2005 – Informatik LIVE (Band 2)
- P-69 Robert Hirschfeld, Ryszard Kowalczyk, Andreas Polze, Matthias Weske (Hrsg.): NODE 2005, GSEM 2005
- P-70 Klaus Turowski, Johannes-Maria Zaha (Hrsg.): Component-oriented Enterprise Application (COAE 2005)
- P-71 Andrew Torda, Stefan Kurz, Matthias Rarey (Hrsg.): German Conference on Bioinformatics 2005
- P-72 Klaus P. Jantke, Klaus-Peter Fähnrich, Wolfgang S. Wittig (Hrsg.): Marktplatz Internet: Von e-Learning bis e-Payment
- P-73 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): "Heute schon das Morgen sehen"
- P-74 Christopher Wolf, Stefan Lucks, Po-Wah Yau (Hrsg.): WEWoRC 2005 – Western European Workshop on Research in Cryptology
- P-75 Jörg Desel, Ulrich Frank (Hrsg.): Enterprise Modelling and Information Systems Architecture
- P-76 Thomas Kirste, Birgitta König-Riess, Key Poustchi, Klaus Turowski (Hrsg.): Mobile Informationssysteme – Potentiale, Hindernisse, Einsatz
- P-77 Jana Dittmann (Hrsg.): SICHERHEIT 2006
- P-78 K.-O. Wenkel, P. Wagner, M. Morgens-tern, K. Luzi, P. Eisermann (Hrsg.): Land- und Ernährungswirtschaft im Wandel
- P-79 Bettina Biel, Matthias Book, Volker Gruhn (Hrsg.): Softwareengineering 2006
- P-80 Mareike Schoop, Christian Huemer, Michael Rebstock, Martin Bichler (Hrsg.): Service-Oriented Electronic Commerce
- P-81 Wolfgang Karl, Jürgen Becker, Karl-Erwin Großpietsch, Christian Hochberger, Erik Maehle (Hrsg.): ARCS'06
- P-82 Heinrich C. Mayr, Ruth Breu (Hrsg.): Modellierung 2006
- P-83 Daniel Huson, Oliver Kohlbacher, Andrei Lupas, Kay Nieselt and Andreas Zell (eds.): German Conference on Bioinformatics
- P-84 Dimitris Karagiannis, Heinrich C. Mayr, (Hrsg.): Information Systems Technology and its Applications
- P-85 Witold Abramowicz, Heinrich C. Mayr, (Hrsg.): Business Information Systems
- P-86 Robert Krimmer (Ed.): Electronic Voting 2006
- P-87 Max Mühlhäuser, Guido Röbling, Ralf Steinmetz (Hrsg.): DELFI 2006: 4. e-Learning Fachtagung Informatik
- P-88 Robert Hirschfeld, Andreas Polze, Ryszard Kowalczyk (Hrsg.): NODE 2006, GSEM 2006
- P-90 Joachim Schelp, Robert Winter, Ulrich Frank, Bodo Rieger, Klaus Turowski (Hrsg.): Integration, Informationslogistik und Architektur
- P-91 Henrik Stormer, Andreas Meier, Michael Schumacher (Eds.): European Conference on eHealth 2006
- P-92 Fernand Feltz, Benoît Otjacques, Andreas Oberweis, Nicolas Poussing (Eds.): AIM 2006
- P-93 Christian Hochberger, Rüdiger Liskowsky (Eds.): INFORMATIK 2006 – Informatik für Menschen, Band 1
- P-94 Christian Hochberger, Rüdiger Liskowsky (Eds.): INFORMATIK 2006 – Informatik für Menschen, Band 2
- P-95 Matthias Weske, Markus Nüttgens (Eds.): EMISA 2005: Methoden, Konzepte und Technologien für die Entwicklung von dienstbasierten Informationssystemen
- P-96 Saartje Brockmans, Jürgen Jung, York Sure (Eds.): Meta-Modelling and Ontologies
- P-97 Oliver Göbel, Dirk Schadt, Sandra Frings, Hardo Hase, Detlef Günther, Jens Nedon (Eds.): IT-Incident Mangament & IT-Forensics – IMF 2006

- P-98 Hans Brandt-Pook, Werner Simonsmeier und Thorsten Spitta (Hrsg.): Beratung in der Softwareentwicklung – Modelle, Methoden, Best Practices
- P-99 Andreas Schwill, Carsten Schulte, Marco Thomas (Hrsg.): Didaktik der Informatik
- P-100 Peter Forbrig, Günter Siegel, Markus Schneider (Hrsg.): HDI 2006: Hochschuldidaktik der Informatik
- P-101 Stefan Böttinger, Ludwig Theuvsen, Susanne Rank, Marlies Morgenstern (Hrsg.): Agrarinformatik im Spannungsfeld zwischen Regionalisierung und globalen Wertschöpfungsketten
- P-102 Otto Spaniol (Eds.): Mobile Services and Personalized Environments
- P-103 Alfons Kemper, Harald Schöning, Thomas Rose, Matthias Jarke, Thomas Seidl, Christoph Quix, Christoph Brochhaus (Hrsg.): Datenbanksysteme in Business, Technologie und Web (BTW 2007)
- P-104 Birgitta König-Ries, Franz Lehner, Rainer Malaka, Can Türker (Hrsg.) MMS 2007: Mobilität und mobile Informationssysteme
- P-105 Wolf-Gideon Bleek, Jörg Raasch, Heinz Züllighoven (Hrsg.) Software Engineering 2007
- P-106 Wolf-Gideon Bleek, Henning Schwentner, Heinz Züllighoven (Hrsg.) Software Engineering 2007 – Beiträge zu den Workshops
- P-107 Heinrich C. Mayr, Dimitris Karagiannis (eds.) Information Systems Technology and its Applications
- P-108 Arslan Brömme, Christoph Busch, Detlef Hühnlein (eds.) BIOSIG 2007: Biometrics and Electronic Signatures
- P-109 Rainer Koschke, Otthein Herzog, Karl-Heinz Rödiger, Marc Ronthaler (Hrsg.) INFORMATIK 2007 Informatik trifft Logistik Band 1
- P-110 Rainer Koschke, Otthein Herzog, Karl-Heinz Rödiger, Marc Ronthaler (Hrsg.) INFORMATIK 2007 Informatik trifft Logistik Band 2
- P-111 Christian Eibl, Johannes Magenheimer, Sigrid Schubert, Martin Wessner (Hrsg.) DeLFI 2007: 5. e-Learning Fachtagung Informatik
- P-112 Sigrid Schubert (Hrsg.) Didaktik der Informatik in Theorie und Praxis
- P-113 Sören Auer, Christian Bizer, Claudia Müller, Anna V. Zhdanova (Eds.) The Social Semantic Web 2007 Proceedings of the 1st Conference on Social Semantic Web (CSSW)
- P-114 Sandra Frings, Oliver Göbel, Detlef Günther, Hardo G. Hase, Jens Nedon, Dirk Schadt, Arslan Brömme (Eds.) IMF2007 IT-incident management & IT-forensics Proceedings of the 3rd International Conference on IT-Incident Management & IT-Forensics
- P-115 Claudia Falter, Alexander Schliep, Joachim Selbig, Martin Vingron and Dirk Walther (Eds.) German conference on bioinformatics GCB 2007
- P-116 Witold Abramowicz, Leszek Maciszek (Eds.) Business Process and Services Computing 1st International Working Conference on Business Process and Services Computing BPSC 2007
- P-117 Ryszard Kowalczyk (Ed.) Grid service engineering and management The 4th International Conference on Grid Service Engineering and Management GSEM 2007
- P-118 Andreas Hein, Wilfried Thoben, Hans-Jürgen Appelrath, Peter Jensch (Eds.) European Conference on ehealth 2007
- P-119 Manfred Reichert, Stefan Strecker, Klaus Turowski (Eds.) Enterprise Modelling and Information Systems Architectures Concepts and Applications
- P-120 Adam Pawlak, Kurt Sandkuhl, Wojciech Cholewa, Leandro Soares Indrusiak (Eds.) Coordination of Collaborative Engineering - State of the Art and Future Challenges
- P-121 Korbinian Herrmann, Bernd Bruegge (Hrsg.) Software Engineering 2008 Fachtagung des GI-Fachbereichs Softwaretechnik
- P-122 Walid Maalej, Bernd Bruegge (Hrsg.) Software Engineering 2008 - Workshopband Fachtagung des GI-Fachbereichs Softwaretechnik

- P-123 Michael H. Breitner, Martin Breunig, Elgar Fleisch, Ley Pousttchi, Klaus Turowski (Hrsg.)
Mobile und Ubiquitäre Informationssysteme – Technologien, Prozesse, Marktfähigkeit
Proceedings zur 3. Konferenz Mobile und Ubiquitäre Informationssysteme (MMS 2008)
- P-124 Wolfgang E. Nagel, Rolf Hoffmann, Andreas Koch (Eds.)
9th Workshop on Parallel Systems and Algorithms (PASA)
Workshop of the GI/ITG Special Interest Groups PARS and PARVA
- P-125 Rolf A.E. Müller, Hans-H. Sundermeier, Ludwig Theuvsen, Stephanie Schütze, Marlies Morgenstern (Hrsg.)
Unternehmens-IT:
Führungsinstrument oder Verwaltungsbürde
Referate der 28. GIL Jahrestagung
- P-126 Rainer Gimnich, Uwe Kaiser, Jochen Quante, Andreas Winter (Hrsg.)
10th Workshop Software Reengineering (WSR 2008)
- P-127 Thomas Kühne, Wolfgang Reisig, Friedrich Steimann (Hrsg.)
Modellierung 2008
- P-128 Ammar Alkassar, Jörg Siekmann (Hrsg.)
Sicherheit 2008
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 4. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)
2.-4. April 2008
Saarbrücken, Germany
- P-129 Wolfgang Hesse, Andreas Oberweis (Eds.)
Sigsand-Europe 2008
Proceedings of the Third AIS SIGSAND European Symposium on Analysis, Design, Use and Societal Impact of Information Systems
- P-130 Paul Müller, Bernhard Neumair, Gabi Dreö Rodosek (Hrsg.)
1. DFN-Forum Kommunikations-technologien Beiträge der Fachtagung
- P-131 Robert Krimmer, Rüdiger Grimm (Eds.)
3rd International Conference on Electronic Voting 2008
Co-organized by Council of Europe, Gesellschaft für Informatik and E-Voting.CC
- P-132 Silke Seehusen, Ulrike Lucke, Stefan Fischer (Hrsg.)
DeLFI 2008:
Die 6. e-Learning Fachtagung Informatik
- P-133 Heinz-Gerd Hegering, Axel Lehmann, Hans Jürgen Ohlbach, Christian Scheideler (Hrsg.)
INFORMATIK 2008
Beherrschbare Systeme – dank Informatik Band 1
- P-134 Heinz-Gerd Hegering, Axel Lehmann, Hans Jürgen Ohlbach, Christian Scheideler (Hrsg.)
INFORMATIK 2008
Beherrschbare Systeme – dank Informatik Band 2
- P-135 Torsten Brinda, Michael Fothe, Peter Hubwieser, Kirsten Schlüter (Hrsg.)
Didaktik der Informatik – Aktuelle Forschungsergebnisse
- P-136 Andreas Beyer, Michael Schroeder (Eds.)
German Conference on Bioinformatics GCB 2008
- P-137 Arslan Brömme, Christoph Busch, Detlef Hühnlein (Eds.)
BIOSIG 2008: Biometrics and Electronic Signatures
- P-138 Barbara Dinter, Robert Winter, Peter Chamoni, Norbert Gronau, Klaus Turowski (Hrsg.)
Synergien durch Integration und Informationslogistik
Proceedings zur DW2008

The titles can be purchased at:

Köllen Druck + Verlag GmbH

Ernst-Robert-Curtius-Str. 14 · D-53117 Bonn

Fax: +49 (0)228/9898222

E-Mail: druckverlag@koellen.de