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 [Ol03]. 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.

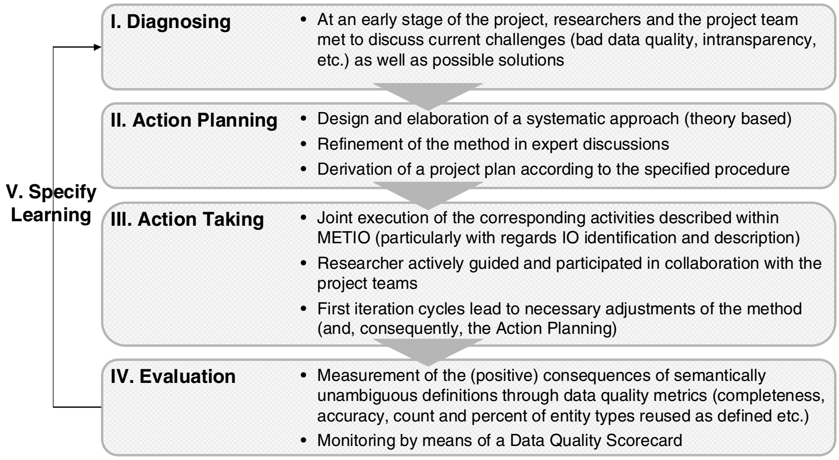

Figure 1: Pursued research process

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 compliance [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].

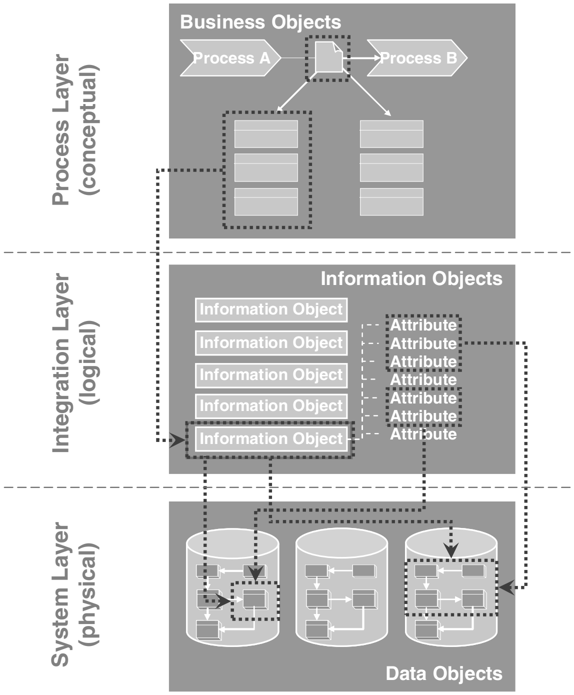

Figure 2: Three-layer differentiation between business, information and data objects

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

FigureFigure 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.

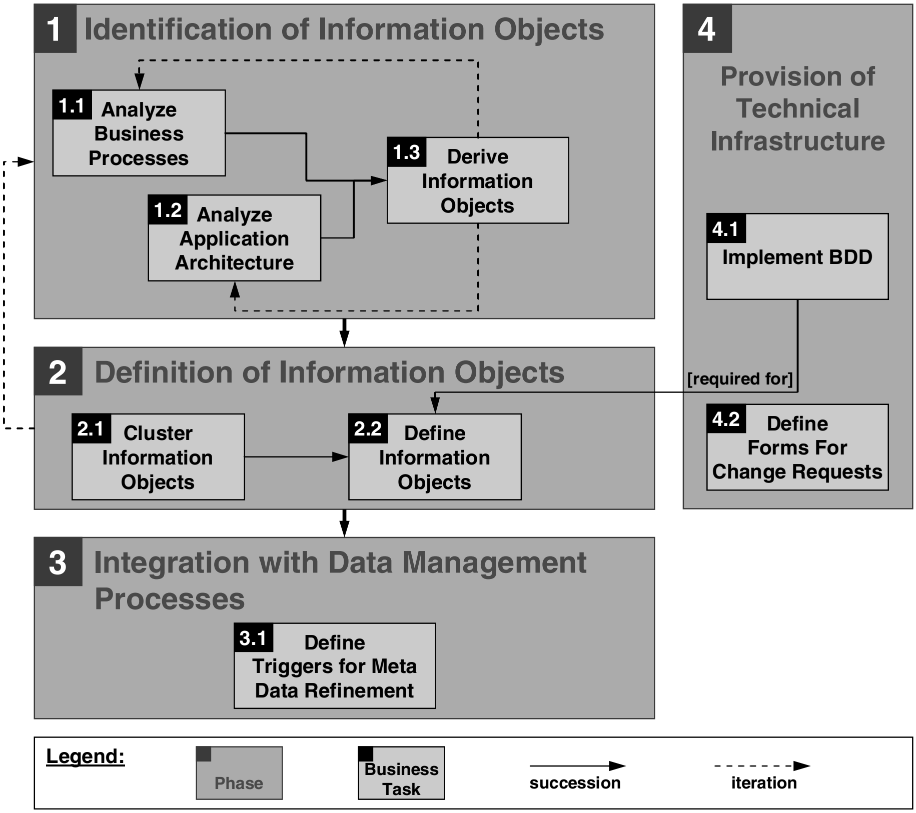

Figure 3: Procedure model of the METIO method

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].

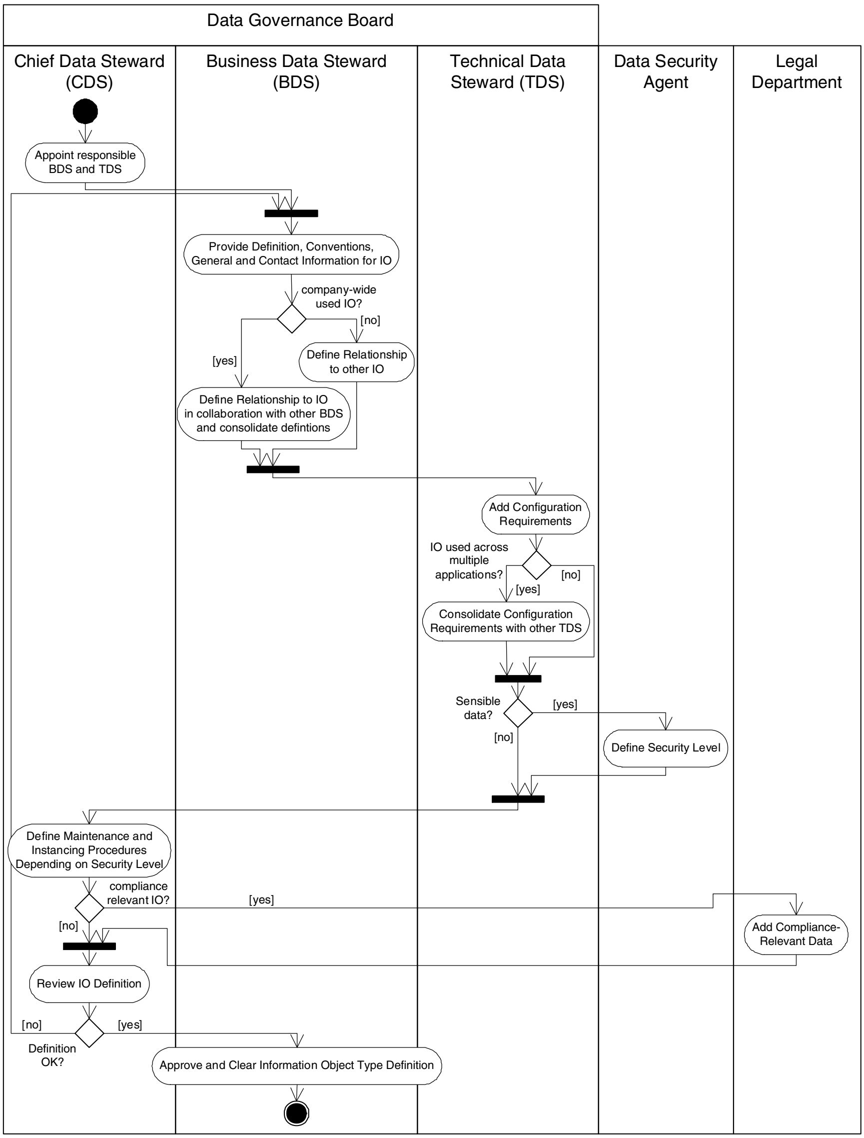

Figure 4: Activity diagram for the definition of information objects

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. Therefor, 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.

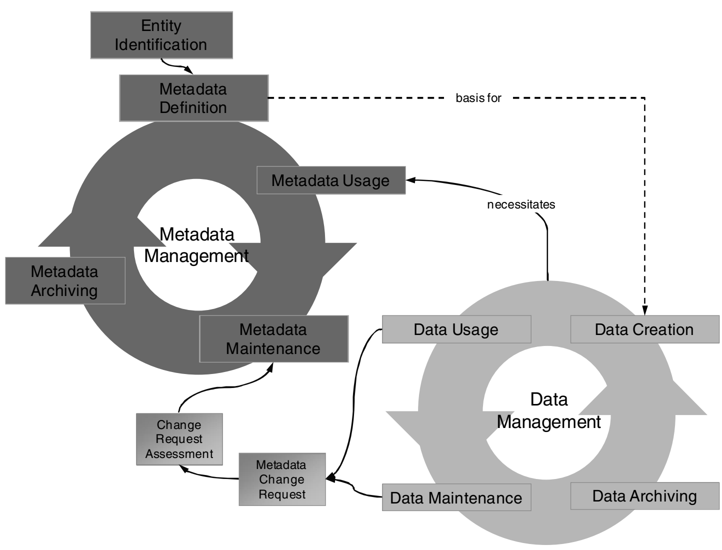

Figure 5: Metadata and data management in a conjoint lifecycle

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.	o Rationale o 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 precisely 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.

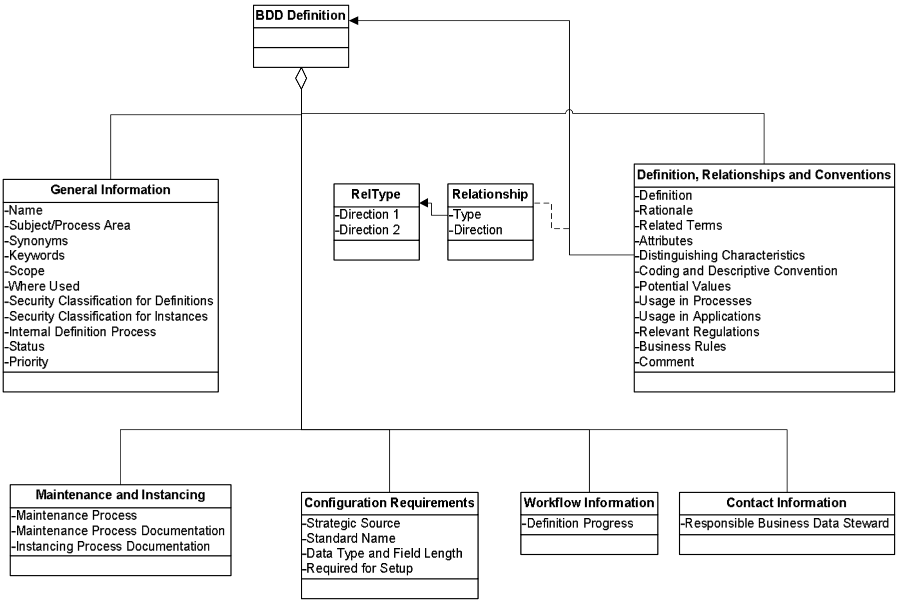

Figure 6: Attributes to be defined for comprehensive IO description in a BDD

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.

[Ol03] 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.