

Informationsmanagement unter Verwendung von freier Software

Alexander Wehrum¹, Claus Mücke²

¹ Connecta AG – Gesellschaft für innovative Informationstechnologie
Rheinstraße 17-21
D-65185 Wiesbaden

² Professur für Biometrie und Populationsgenetik
Justus-Liebig-Universität
Heinrich-Buff-Ring 26-32
D-35392 Giessen
info@connecta.ag
claus.mueckschel@agrار.uni-giessen.de

Abstract: Heterogene und stetig wachsende Datenmengen können nur unter Verwendung moderner Informationstechnologien verwaltet werden. Content- und Dokumentenmanagement-Lösungen haben sich als ein integraler Bestandteil des Informationsmanagements für Forschung und Praxis etabliert. Der Beitrag stellt diese beiden Softwarelösungen kurz vor und beschreibt Kriterien, die bei Evaluierung von freier Software in diesem Bereich zu beachten sind.

1 Einleitung

Moderne Informations- und Kommunikationstechniken sind unlängst zu einer maßgeblichen Erfolgskomponente für Projekte in Wirtschaft und Forschung geworden. Besonders im Rahmen von Forschungsprojekten werden meist große und heterogene Datenmengen erzeugt, die oftmals auch nach Zeiträumen von über 10 Jahren in einem angemessenen Zeitraum auffindbar und reproduzierbar sein müssen. Das gezielte Management von Informationen soll sicherstellen, dass berechnigte Personen relevante Daten in einem angemessenen Zeitraum zur Verfügung gestellt bekommen. Langfristig ausgelegte und sinnvoll konzipierte Informationssysteme sind demnach Grundlage für den Projekterfolg. Bei zahlreichen Projekten in der freien Wirtschaft sowie im Rahmen einer Kooperation mit dem Daten- und Informationsmanagement des DFG-Sonderforschungsbereichs 299 (Justus-Liebig-Universität Gießen) konnten die Autoren Praxiserfahrungen im Umgang mit Informationssystemen, insbesondere auf Basis von Open-Source (OS) Software, sammeln. Um strukturierte und weniger strukturierte Informationen, wie sie in Dokumenten vorliegen, verfügbar und damit nutzbar zu machen, werden unter anderem Content-Management-Systeme (CMS) und spezielle Komponenten, wie Dokumenten-Management-Systeme (DMS), eingesetzt, für die mittlerweile leistungsstarke OS-Lösungen existieren.

2 Content-Management-Systeme

Content-Management-Systeme wurden ursprünglich meist zur Unterstützung redaktioneller Prozesse bei der Inhaltserstellung von Websites eingesetzt.

Mittlerweile sind die Grenzen zwischen Website-Inhaltsverwaltung und einer gezielten Bereitstellung von Informationen im Intra- oder Extranet allerdings unscharf geworden. Ein CMS ist nunmehr auch ein wesentlicher Baustein des Informationsmanagements. Ein spezifisches Merkmal aller CMS ist die Trennung von Inhalt und der zugehörigen Darstellung. Die Systeme unterstützen die Informationseingabe durch definierbare Workflows. Eingabe, Abruf und Verwaltung der Informationen können dabei browserbasiert, d.h. weitgehend plattformunabhängig vorgenommen werden. Darüber hinaus wird es durch die Trennung von Inhalt und Darstellung möglich, Inhalte in unterschiedlichen Kontexten, Kombinationen, Medien und Formaten zu präsentieren. Hierdurch wird eine medienneutrale Datenhaltung im Sinne des „Crossmedia Publishing“ erreicht. Dies ermöglicht nicht nur eine flexible Handhabung im Datenmanagement, sondern darüber hinaus auch maximale Transparenz bei Organisation, Verwaltung, Pflege und Export der Daten sowie die Vermeidung von Fehlern und von Redundanzen bei der Datenhaltung.

3 Dokumenten-Management-Systeme

Mit Dokumenten-Management-Systemen werden große und heterogene Dokumente zentral verwaltet. Bei DM-Systemen liegt der Fokus weniger auf der redaktionellen Inhaltseingabe und der Trennung von Inhalt und Darstellung, wie bei CM-Systemen, sondern vielmehr auf der Logik der effektiven Dokumentenverwaltung. Dokumente, die mittels eines DMS verwaltet werden, werden dabei um spezifische Informationen, die so genannten Metadaten, erweitert [MSK05]. Typische Metadaten bestehen meist aus Merkmalen wie einer eindeutigen Schlüsselnummer sowie aus Regelattributen, die beispielsweise Auskunft über Lebensdauer oder Zugriffsrechte geben. Darüber hinaus wird der Inhalt des Dokuments kategorisiert, inhaltlich beschrieben und in Relation zu anderen Dokumenten gesetzt. Mit den generierten Metadaten kann mittels einer strukturierten Suche ein entsprechendes Dokument recherchiert werden [MSK05].

4 Besonderheiten freier Software für das Informationsmanagement

Mittlerweile gibt es in Teilbereichen des Informationsmanagements ausgereifte OS-Lösungen. Um aus der Vielzahl von freien und proprietären Alternativen eine für die eigenen Bedürfnisse optimale Informationsmanagementlösung zu finden, ist ein systematischer Evaluierungsprozess unumgänglich. Hierbei werden in Abhängigkeit von der Menge der zu evaluierenden Lösungen zunächst in einer Grobauswahl durch eine Definition von Ausschluss-Kriterien gänzlich ungeeignete Lösungen - z.B. wegen fehlender Grundfunktionalitäten - ausgeschlossen. Parallel dazu werden einzelne qualitative und quantitative Evaluierungskriterien festgelegt, etwa die Gesamtkosten und/ oder die Benutzerfreundlichkeit und wegen ihrer unterschiedlich großen Relevanz für den Projekterfolg entsprechend ihrer Bedeutung gewichtet.

Mit Hilfe der gewichteten Auswahlkriterien wird die optimale Softwarelösung bestimmt. Dabei ist zu beachten, dass nur solche Kriterien festgelegt werden, die später operational formuliert und mit einem angemessenen Aufwand messbar sind. Grundsätzlich müssen Funktionalität, Anpassungsmöglichkeiten, Modularität, Benutzerfreundlichkeit, Charakteristika des Herstellers, Sicherheit und schließlich die Total Cost of Ownership (TCO) beim Auswahlprozess berücksichtigt werden. Werden proprietäre und freie Softwarelösungen im Rahmen des Evaluierungsprozesses verglichen, sind Open-Source-spezifische Kriterien zu berücksichtigen. Ein Open-Source-spezifisches Evaluierungskriterium ist die weitgehende Unabhängigkeit von Herstellern, die proprietäre Lösungen i.d.R. nicht aufweisen. Dies kann gerade bei langfristig ausgelegten Forschungsprojekten, wie dem im Jahr 1997 gestarteten SFB 299, von großem Vorteil sein. Eine mögliche Folge des Einsatzes proprietärer Software könnte sein, dass man zu kosten- und zeitintensiven Migrationen auf ein neues Produkt gezwungen wird, da der Support eingestellt wird oder ein Hersteller gänzlich vom Markt verschwindet. Hier bieten OS-Lösungen oftmals Vorteile, da man durch die Offenheit des Quellcodes und die fehlende Notwendigkeit von bilateralen Verträgen mit Herstellern unabhängiger wird. Dieser Umstand führt zudem zu einem erhöhten Wettbewerb und damit meist zu geringen indirekten Transaktionskosten im Rahmen nötiger Anpassungs- und Erweiterungsarbeiten. Wegen dabei entfallender Lizenzkosten können OS-Lösungen auch bei den TCO, also der Summe aus direkten und indirekten Kosten, im Vorteil sein [WEM04]. Allerdings sind OS-Lösungen nicht per se kostengünstiger als proprietäre Lösungen. Gerade der erhöhte Wettbewerb durch immer leistungsfähigere OS-Lösungen setzt die etablierten Softwarehersteller zunehmend unter Preisdruck. Bei Kostensenkungen auf Herstellerseite sinken die Vorteile auf OS-Seite entsprechend. Außerdem neigen Entscheider oft dazu, die indirekten Kosten einer Softwareinvestition zu vernachlässigen. Hier kann eine OS-Lösung durchaus im Nachteil sein. Zwar sind die indirekten Kosten wegen der Open-Source-spezifischen Modularität der Software, der Herstellerunabhängigkeit sowie dem offenen Quellcode oftmals niedriger, allerdings muss man diese Vorteile in Bezug zu den eigenen Anforderungen an eine Software setzen. In der betriebswirtschaftlichen Praxis ist die Möglichkeit, den Quellcode zu modifizieren oder die Modularität oftmals nur von geringer Bedeutung, was den Kostenvorteil in dieser Hinsicht relativiert. Im Rahmen des Daten- und Informationsmanagements in Forschungsprojekten sind diese Eigenschaften hingegen oft von großer Bedeutung. Zum einen ist die Modularität und Anpassbarkeit dort wegen der heterogenen Datenquellen und des ständigen Anpassungsdruckes sehr wichtig und zum anderen ist meist projektinterne Fachkompetenz vorhanden, die Anpassungen bei OS-Lösungen kostengünstig realisieren zu können. Der Kostenvorteil freier Software hinsichtlich der direkten Kosten ist außerdem besonders ausgeprägt bei Serverbetriebssystemen, Datenbanken oder ähnlich zentraler Komponenten, da deren Lizenzkosten oft durch Client Access Licenses (CLAs) mit der Anzahl der angeschlossenen Nutzer steigen (Abb. 1).

Insgesamt ist der Kostenvorteil im Bereich des Informationsmanagements im universitärem Umfeld nach Erfahrung der Autoren meist größtenteils auf Seite der freien Systeme. Ein Content-Management-System, wie beispielsweise TYPO3, ist ein Beispiel für eine ausgereifte und zudem kostengünstige Open-Source-Softwarelösung für vielfältige Aufgabenstellungen im Rahmen des Informationsmanagements [WEM04]. Es gibt allerdings auch eine Reihe möglicher Nachteile von OS-Software, wie zum Beispiel die teilweise ungeklärte Patentrechtsproblematik und die fehlende Gewährleistung und Haftung eines Herstellers. In jedem Fall sollte deswegen darauf geachtet werden, dass die Nutzergemeinschaft groß und aktiv ist, denn nur so kann sichergestellt werden, dass bei möglichen Problemen Lösungen zeitnah zur Verfügung stehen.

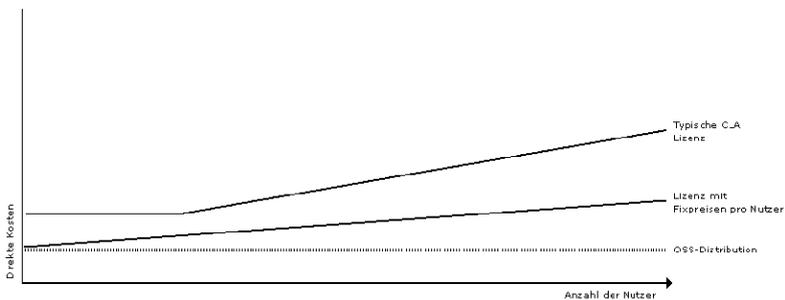


Abbildung 1: Typischer Verlauf der direkten Kosten bei steigender Anzahl der Softwarenutzer [Wi05].

Außerdem sollte eine weit verbreitete Lizenzart, wie die GPL (General Public License), vorgezogen werden. Allerdings muss einem Open-Source-Modell sowohl die Software als auch der zugrunde liegende Quellcode frei verfügbar sein, weshalb auch der Entwickler zumeist verpflichtet wird, den eigens modifizierten oder erweiterten Quellcode wieder der Öffentlichkeit zugänglich zu machen. Dies kann dann ins Gewicht fallen, wenn mit der eingesetzten Software ein Wettbewerbsvorteil erreicht werden soll. Dieser liegt vor, wenn ein bestimmter Faktor - in unserem Beispiel die Software - wertvoll, selten, nicht oder nicht vollständig imitierbar bzw. nicht ersetzbar ist [Ba91]. Der Wettbewerbsvorteil kann folglich wegen der Verpflichtung zur Veröffentlichung unter Umständen mit OS-Software nicht erreicht werden.

5 Fazit und Ausblick

Open-Source-Lösungen im Bereich des Informationsmanagements können äußerst sinnvolle Werkzeuge sein. Besonders im universitären Umfeld wurden überwiegend positive Erfahrungen mit dem OS-Ansatz gemacht. Gegenwärtig finden Überlegungen statt, das Open-Source CMS TYPO3 und ein zugehöriges Metadatenmanagement-Modul im Rahmen des Daten- und Informationsmanagement-Projekts des Sonderforschungsbereichs 299 einzusetzen. Das dazugehörige Metadatenmanagement-Modul, das ähnlich wie ein DMS die Verwaltung von Dokumenten und der dazugehörigen Metadaten übernimmt, erfüllt die wesentlichen Kriterien einer Open-Source-Software. Jedoch kann der Entwicklungsstand aktuell nur als rudimentär bezeichnet werden. Hier finden erste Tests statt, die zu Überlegungen geführt haben, das Metadatenmanagement der Software TYPO3 in Eigenentwicklung auszubauen und die Software so auf die eigenen Bedürfnisse anzupassen.

Literaturverzeichnis

- [Ba91] Barney, J.: Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 1991; S. 99-120.
- [MSK05] Mückschel, C., Schachtel, G.A., Köhler, W.: Management of metadata in an interdisciplinary environmental research project. - *Proceedings of the EFITA/ WCCA 2005 Conference*, Vila Real, Portugal, 2005; S. 1434-1441.
- [Wi05] Wichmann, T.: *Linux und Open-Source-Strategien*. Springer, Berlin: 36, 2005.
- [WEM04] Wehrum, A.; Eulberg, P.; Mückschel, C.: Paradigmen des Open-Source Konzeptes am Beispiel der Software TYPO3. *Zeitschrift für Agrarinformatik* 4, 2004; S. 104-111.