

MCI-DL: Eine digitale Bibliothek für Mensch-Computer-Interaktion

Philipp Bohnenstengel, Martin Christof Kindsmüller, Michael Herczeg

Institut für Multimediale und Interaktive Systeme (IMIS), Universität zu Lübeck

Zusammenfassung

In diesem Beitrag wird die derzeit im Aufbau befindliche digitale Bibliothek des Fachbereichs MCI der GI vorgestellt. Inhalt der MCI-DL werden zunächst die Beiträge der Konferenzreihe Mensch & Computer und der Vorgängerkonferenz Software-Ergonomie seit 1983 sein. Es wird auf das zugrunde liegende Framework DSpace eingegangen und anhand von Beispielen gezeigt, wie die MCI-DL ganzheitliche Arbeitsabläufe unterstützen kann, welche über die Nutzung des Systems selbst hinausgehen.

1 Einleitung

Literatur-Recherche im World Wide Web ist heutzutage eine wichtige Informationsquelle im Rahmen der wissenschaftlichen Arbeit. Dazu beigetragen haben große Fachgesellschaften und Fachverlage, die – meist gegen Gebühr – die Publikationen ihrer Autoren im Web zugänglich machen. Auch die sich verbreitende Philosophie des Open Access, im Namen derer kostenfreier Zugang zu wissenschaftlichen Publikationen und liberale Nutzungsrechte gewährt werden (vgl. Suber, 2002; Bullinger et al., 2003), hat ihren Anteil daran. Eine der tragenden Säulen dieser Entwicklung sind digitale Bibliotheken (Digital Libraries), die wie ihre Gegenstücke in der physikalischen Welt einen Bestand von Publikationen sammeln, diesen pflegen und strukturierten Zugriff auf diesen ermöglichen. Neben dem offensichtlichen Vorteil, dass sich digitale Dokumente ohne Aufwand beliebig oft vervielfältigen lassen, kann eine digitale Bibliothek im Gegensatz zu einer physikalischen ohne großen Zusatzaufwand fast beliebig viele Perspektiven auf ihren Bestand bieten und semantische Beziehungen zwischen verschiedenen Dokumenten ausnutzen.

Der Aufbau einer solchen digitalen Bibliothek für das Gebiet der Mensch-Computer-Interaktion (MCI-DL), insbesondere für Inhalte der Konferenz „Mensch & Computer“ (M&C) und der Vorgängerkonferenz „Software-Ergonomie“ (SE), wurde vom Fachbereich MCI der Gesellschaft für Informatik angestoßen, welche – im Gegensatz zu ihren US-amerikanischen bzw. internationalen Pendanten, der ACM und dem IEEE – noch keine eigene digitale Bibliothek betreibt. Ziel ist es, die Konferenzbände, die derzeit zur Konferenz nur als

Papierversion erscheinen und erst ein halbes Jahr später als statische Website veröffentlicht werden, direkt vom Konferenzmanagementsystem in die MCI-DL zu transferieren. So können bereits auf der Konferenz alle Teilnehmer online auf die Beiträge zugreifen. Im Gegensatz zu einer statischen Website kann bei einer digitalen Bibliothek leichter sichergestellt werden, dass auch bei stetig wachsender Dokumentanzahl die Übersichtlichkeit erhalten bleibt. Gleichzeitig kann über das Bereitstellen von Such- und Filtermöglichkeiten ein deutlicher Mehrwert gegenüber einer statischen Website geschaffen werden. Die Grundlagen für die MCI-DL wurden in Bohnenstengel (2010) beschrieben. Aufbauend auf einer umfassenden Aufgaben-, Benutzer- und Kontextanalyse sowie einer umfangreichen Sichtung und Bewertung von Digital-Library-Frameworks wurde die MCI-DL konzipiert und die zentralen Use-Cases prototypisch unterstützt. Der vorliegende Beitrag orientiert sich im Aufbau an dieser Arbeit, setzt aber als „ausstellungsbegleitender Kurzbeitrag“ abweichende Schwerpunkte, um die Systemdemonstration theoretisch zu unterfüttern und zu motivieren.

2 Technische Basis: DSpace

Die Basisfunktionen einer digitalen Bibliothek sind klar abgesteckt: Die Inhalte müssen sich strukturiert präsentieren lassen, so dass sie per zielgerichteter Navigation erreichbar sind und es bedarf einer Suchfunktion, die sowohl Inhalte (Abstract, Volltext) als auch beschreibende Metadaten erfasst. Zusätzlich ist ein Unterbau nötig, der das Erzeugen und Verwalten der dafür notwendigen Datenstrukturen ermöglicht. Da es speziell auf diese Anforderungen zugeschnittene Frameworks gibt, die zudem im Bereich der digitalen Bibliotheken etablierte Standards implementieren, wäre eine Eigenentwicklung auf Basis eines allgemeinen Content-Management-Systems wie z.B. Drupal nicht sinnvoll gewesen. Von den Frameworks für digitale Bibliotheken haben sich in den letzten Jahren vor allem DSpace¹ und EPrints¹ breit etabliert, wobei DSpace nach unabhängigen Zählungen² eine deutlich größere installierte Basis aufweist. Da sich die Frameworks in ihrem Funktionsumfang nicht wesentlich unterscheiden, fiel die Wahl aufgrund der größeren Verbreitung und der als aktiver eingeschätzten Entwicklergemeinschaft auf DSpace.

DSpace ist eine Open-Source-Software, die ihrerseits unter anderem auf Java-EE (Apache Tomcat) und PostgreSQL aufbaut und auf eine Kooperation des MIT und der HP-Labs zurückgeht. DSpace ist auf den Betrieb sogenannter *Institutional Repositories* zugeschnitten und bietet einen großen Umfang an grundlegender Funktionalität für den Betrieb einer digitalen Bibliothek. Dies umfasst sowohl den Bereich der Verwaltung und Organisation von Dokumenten, als auch den des öffentlichen Zugriffs auf diese. Darüber hinaus werden eine Reihe von etablierten Standards wie *Dublin Core*, *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) und *Open Search* unterstützt.

¹ DSpace: <http://www.dspace.org> Stand; EPrints: <http://www.eprints.org/software>

² University of Nottingham: <http://www.openoar.org>

3 Die MCI-DL als Bestandteil komplexer Aufgaben

Eine digitale Bibliothek steht nicht für sich allein. Sie ist innerhalb des wissenschaftlichen Betriebs Teil eines „Ökosystems“ von Software, Diensten und Arbeitsmethoden, die sich gegenseitig ergänzen und miteinander verwoben sind (vgl. Abbildung 1). Diese komplexen Zusammenhänge wurden im Rahmen der Aufgaben- und Benutzeranalyse durch strukturierte Interviews mit Wissenschaftlern und Studierenden untersucht, in denen Arbeitsabläufe und Gewohnheiten bei der Literaturrecherche im WWW, dabei auftretende Probleme und denkbare Erweiterungen bestehender Systeme zur Sprache kamen. Auf diese Weise wurden Erkenntnisse zu ganzheitlichen Arbeitsabläufen gewonnen. Diese bilden die Ausgangsbasis für die Konzeption der MCI-DL.

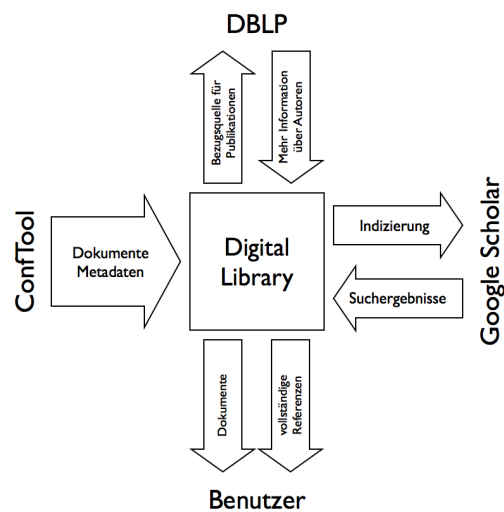


Abbildung 1: Die MCI-DL als Teil eines Ökosystems

Für die Organisation der M&C wird in der Regel die Software ConfTool³ verwendet. Hier werden im Rahmen des Konferenzmanagements ein Großteil der für die MCI-DL relevanten Metadaten und auch die PDF-Versionen aller Beiträge erfasst. Für den Import in die MCI-DL wurde eine Schnittstelle zur direkten Datenübernahme einer mit ConfTool organisierten Konferenz entwickelt. Ein bedeutender Teil wissenschaftlicher Arbeit ist es, auf den Erkenntnissen anderer aufzubauen. Das korrekte Zitieren von Quellen wird durch die MCI-DL unterstützt, indem zu jedem Dokument vollständige Referenzen dynamisch aus den Metadaten erzeugt werden. Diese sind in verschiedenen Formaten wie BibTeX oder dem im Bereich der HCI-Forschung sehr verbreiteten APA-Style verfügbar und können so einfach weiterverarbeitet werden. Da eine digitale Bibliothek – ganz gleich, wie umfangreich sie auch sein mag – immer nur über einen beschränkten Bestand verfügen kann, ist es wichtig, den Nut-

³ <http://www.conftool.net>

zern Verknüpfung zu weiterführenden Diensten jenseits des eigenen Bestands zu bieten. So wird beispielsweise neben der Möglichkeit, alle Artikel eines Autors in der MCI-DL anzuzeigen, ein Hyperlink zu DBLP (Ley, 1997) angeboten, wo i.d.R. umfassende Information zu weiteren Arbeiten des Autors verfügbar sind. Andererseits muss auch Zugriff von außen ermöglicht werden. Hier wird insbesondere eine Indizierung durch Dienste wie Google Scholar und CiteSeerX ermöglicht, da diesen in der Forschung und Lehre in zunehmendem Maße Vorzug vor dem direkten Zugriff auf die Anbieterrepositories gegeben wird. Auch die Verknüpfung mit den internationalen Standardnummernsystemen für Publikationen ISBN und ISSN sowie DOI werden unterstützt. Die MCI-DL soll als stark fachbezogene digitale Bibliothek auch mit einem Fachindex ausgestattet werden, der zusammen mit weiteren semantischen Markups erlauben soll, zielgerichtet und effizient fachliche Inhalte mit spezifischen Perspektiven zu erschließen. Die Verschlagwortung soll an übliche Standards (z.B. ACM) anknüpfen, diese aber im Fachgebiet MCI verfeinern. Umgekehrt kann diese Verschlagwortung die *Call for Papers* für anstehende Konferenzen und Workshops unterstützen und die Entwicklung des Fachgebiets durch eine Bestandsaufnahme und Strukturierung fördern.

4 Ausblick

Die MCI-DL ist in Grundzügen realisiert, wobei die Beiträge der M&C-Konferenzen seit 2001 bereits integriert sind. Da die Integration der Beiträge der SE-Konferenzen und die Arbeiten an der Benutzungsoberfläche bis zur M&C 2011 weitestgehend abgeschlossen sein werden, soll eine erste öffentliche Version dort präsentiert werden. Mit diesem Schritt wird begonnen die MCI-DL – im Sinne eines benutzerzentrierten Entwicklungsprozesses – in der gesamten Fachcommunity zu erproben, zu evaluieren und zu optimieren. Weiterhin müssen effiziente Arbeitsabläufe für die Aufnahme zukünftiger Inhalte etabliert werden. Eine zukünftige Version der MCI-DL könnte auch kollaborative Aspekte beinhalten – beispielsweise zur kollaborativen Generierung des im letzten Abschnitt beschriebenen Fachindex.

Literaturverzeichnis

- Bohnenstengel, P. (2010). *Aufbau einer Digital Library für Mensch-Computer-Interaktion*. Bachelorarbeit am Institut für Multimediale und Interaktive Systeme (IMIS) der Universität zu Lübeck.
- Bullinger, H.-J., Einhüpl, K. M., Gaetgens, P. et al. (2003). Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. *Open Access to Knowledge in the Sciences and Humanities; 20.–22.10.2003, Berlin*.
- Ley, M. (1997). Die Trierer Informatik-Bibliographie DBLP. In Jarke, M., Pasedach, K. & Pohl, K. (Hrsg.): *Informatik '97, Informatik als Innovationsmotor*. Berlin: Springer, S. 257-266.
- Suber, P. (2002). Open access to the scientific journal literature. *Journal of Biology I* (1), S. 3.1-3.3.

Kontaktinformation

Martin Christof Kindsmüller, mck@imis.uni-luebeck.de