

Einführung eines einheitlichen Systems für die Universitätsbibliographie

Open Access Server und Hochschulschriften an der Universität Regensburg

Gernot Deinzer und Rafael Ball

■ Bislang gab es an der Universität Regensburg zwei unterschiedliche Systeme zum elektronischen Veröffentlichen von Forschungsergebnissen. Aufgrund von dadurch entstandenen Unsicherheiten seitens der Bibliothekskunden und einem organisatorischen Mehraufwand erfolgte im Oktober 2009 die Umstellung auf ein einheitliches System, das internationale Standards erfüllt. Dadurch ergab sich eine deutliche Serviceverbesserung für die Wissenschaftler der Universität Regensburg.

Ausgangslage an der Universität

Mitarbeiter der Universität konnten bereits seit 1999 Hochschulschriften auf einem Server der Universitätsbibliothek online veröffentlichen. Technisch realisiert war dieser Vorgänger des jetzigen Systems als reiner Volltextserver, der vor allem für Dissertationen und die Schriftenreihe „Regensburger Diskussionsbeiträge zur Wirtschaftswissenschaft“ sehr gut angenommen wurde. Dieser Server wurde mit der Software OPUS¹ betrieben.

Im Jahr 2006 startete die Universitätsbibliothek Regensburg mit einem institutionellen Repositorium (Publikationsserver) als zusätzliche Dienstleistung. Auf diesem Server konnten Angehörige der Universität ihre Publikationen auflisten und verwalten. Möglich war dabei auch das Einbringen von Volltexten, die im Sinne von Open Access frei im Internet zur Verfügung gestellt wurden. Dazu wurde die Software EPrints² eingesetzt.

Durch den Betrieb unterschiedlicher Systeme ergaben sich bei einigen Kunden der Universitätsbibliothek Unsicherheiten. So konnten Dokumente mit freier Zugänglichkeit in beiden oder nur in einem der beiden Systeme gefunden werden. Dem Kunden war der Unterschied zwischen den Systemen häufig nicht bewusst. Zusätzlich sah sich die Universitätsbibliothek vor dem Problem, zwei Server mit unterschied-

licher Software zu betreiben und Mitarbeiter in zwei unterschiedliche Systeme einzuarbeiten. Deshalb lag die Entscheidung der Universitätsbibliothek nahe, die unterschiedlichen Systeme in eine einheitliche Server- und Softwarelösung zu migrieren.

Anforderungen

Das neue System sollte als institutionelles Repositorium dienen. Dabei sollte es für die Wissenschaftler³ der Universität Regensburg möglich sein, alle ihre Veröffentlichungen in diesem System nachzuweisen. Hier bestand der Konflikt, dass Wissenschaftler einerseits eine komplette Liste ihrer Publikationen einbringen möchten, sie aber andererseits nicht das Recht besitzen, alle ihre Veröffentlichungen frei zugänglich auf einem Server der Universitätsbibliothek zusätzlich zur Verlagspublikation zu veröffentlichen. Ein reiner Open Access Server stellte damit keine ideale Lösung dar. Erforderlich war die Möglichkeit, sowohl rein bibliographische Daten als auch Volltexte in das System einzubringen. Zusätzlich wünschenswert war dabei die Möglichkeit, eine sogenannte Embargoperiode für die Freishaltung von Volltexten zu setzen. Diese werden somit erst nach Erreichen eines definierten Datums frei zugänglich gemacht, da bei manchen Verlagen die Selbstarchivierung erst nach dem Ende einer Sperrfrist möglich ist. Durch das Einrichten von Embargoperioden können Wissenschaftler der Universität bereits beim bibliographischen Eintrag den Volltext auf den Server stellen und brauchen nach dem Ablauf der Sperrfrist keine weiteren Aktionen durchzuführen.

Idealerweise sollten die Dokumente sowohl als Primärveröffentlichung der Universität, als auch parallel zu einer Verlagspublikation öffentlich zugänglich gemacht werden können. Dabei war vor allem bei den Primärpublikationen darauf zu achten, dass die dauerhafte Zitierfähigkeit gewährleistet ist. Zu

diesem Zweck war die Vergabe einer URN (Uniform Resource Name)⁴ zwingend erforderlich.

Eine weitere Anforderung war das Erstellen von Publikationslisten. Wissenschaftler wollen für unterschiedliche Einsatzmöglichkeiten Publikationslisten zusammenstellen und diese nachnutzen. So sollte es möglich sein, eine Liste nach bestimmten Kriterien zu generieren und entweder in ein Literaturverwaltungsprogramm oder auf verschiedene Internetseiten (z. B. der Fakultät, der Arbeitsgruppe, eines Forschungsprojekts) einzubringen. Wichtig war hierbei, dass jedem Mitarbeiter der Universität eine eindeutige Identifikationsnummer zugeordnet wird. Ziel ist hier, dass Mitarbeiter der Universität ihre Veröffentlichungen an einer zentralen Stelle online pflegen können und redundante Mehrfacharbeit vermieden wird. In der Zukunft soll das System auch als vollständige Universitätsbibliographie dienen.

Darüber hinaus sollte der Publikationsserver keine Stand-alone-Lösung darstellen, sondern mit weiteren, vor allem fachlichen, Repositorien vernetzt werden. Dabei war es der Universitätsbibliothek Regensburg wichtig, dass der Wissenschaftler die freie Wahl hat, ob er seine Einträge in ein fachliches Repositorium oder in das institutionelle Repositorium einbringen möchte. Das institutionelle Repositorium musste also sowohl Import- als auch Exportfunktionen besitzen.

Auch für weitere Austauschfunktionen musste das System eine standardisierte OAI-Schnittstelle⁵ aufweisen. Diese sollte einfach erweiterbar sein, damit spezielle Dienste, wie die Vergabe einer URN, möglich sind.

Entscheidung für ein geeignetes System

Bei der Auswahl eines Systems, das diesem Anforderungsprofil entspricht, beschränk-

1 Siehe <http://opusdev.bsz-bw.de/trac>

2 Siehe <http://www.eprints.org/>

3 Wenn im Folgenden die verallgemeinernde männliche Form für die Bezeichnung von Personengruppen verwendet wird, sind sowohl weibliche als auch männliche Personen gemeint.

4 Siehe <http://www.persistent-identifier.de/>

5 Siehe <http://www.openarchives.org/>.



Abbildung 1: Der Rektor der Universität Regensburg Prof. Dr. Thomas Strothotte weicht zusammen mit dem Direktor der Universitätsbibliothek Regensburg Dr. Rafael Ball den Publikationsserver der Universität Regensburg offiziell ein (Universitätsbibliothek Regensburg, 21. Oktober 2009).

ten wir uns auf die beiden Softwarelösungen OPUS und EPrints, da diese bereits an der Universitätsbibliothek Regensburg im Einsatz waren und somit sehr viel Erfahrung mit diesen Softwarepaketen vorhanden war. Diese sind zusammen mit DSpace⁶ die beiden weltweit am meisten eingesetzten Systeme. Laut dem Registry of Open Access Repositories⁷ werden 649 Repositorien mit DSpace, 336 mit EPrints und 40 mit OPUS betrieben. Bei OpenDOAR⁸ sind 592 DSpace-, 280 EPrints- und 54 OPUS-Installationen registriert. Anhand der genannten Anforderungen wurden die beiden bisher bestehenden Systeme analysiert und das für unsere Zwecke geeignetere System bestimmt. Die Universitätsbibliothek Regensburg hat sich entschieden, die Software EPrints für den neuen Service einzusetzen. Die Gründe hierfür waren:

Das System sollte eine einfache Benutzerverwaltung mit sich bringen. Dabei ist es wünschenswert, dass dieses mit dem bereits bestehenden System der Universität Regensburg verbunden ist. Eine Anbindung mittels Shibboleth⁹ oder Lightweight Directory Access Protocol (LDAP)¹⁰ ist hier notwendig. Sowohl Shibboleth als auch LDAP lassen sich mit EPrints verwirklichen. Die Rechenzentrumskennung soll dabei als eindeutige ID, sowohl als Benutzerkennung als auch zur eindeutigen Bezeichnung für Autoren und Herausgeber dienen. Wissenschaftler, die nicht der Universität Regensburg angehören, können ihre E-Mail-Adresse angeben, die somit zur eindeutigen Identifikation dienen kann. Damit diese auch von

Jedem ohne große Suche eingegeben werden kann, wird bei der Eingabe von Autoren bzw. Herausgebern bereits nach Einträgen von Personen mit dem Namensbeginn gesucht und es werden Vorschläge angezeigt, welche durch einen einfachen Klick übernommen werden können (Technik: AJAX¹¹).

Auch die Struktur der Universität muss flexibel darstellbar sein. Gerade bei Änderungen der Fakultätsstruktur, neuen Lehrstühlen und Arbeitsgruppen oder der Versetzung eines Professors in den Ruhestand, soll die neue Struktur richtig abgebildet werden. Für die Mitarbeiter der Universität sollen keine Änderungen nötig und die Strukturveränderung für die Bibliothek mit einem einmaligen, geringen Aufwand umzusetzen sein. Durch die Verwendung von EPrints ist dies komfortabel möglich.

Wichtig ist die möglichst einfache Eingabe aller relevanten Daten für den Benutzer. Neben der Autovervollständigung der Personen gibt es auch Vorschläge für die Titel von Zeitschriften, Projekten, Workshops, Schriftenreihen usw. Im Falle der Zeitschriften wird dabei eine bestehende Datenbank zum Vorschlag von Zeitschriftentiteln verwendet. Es werden nicht nur der Zeitschriftentitel ergänzt, sondern zusätzlich auch die ISSN und der Verlag eingefügt. Gerade diese Daten werden erfahrungsgemäß selten eingegeben. Zusätzlich werden noch Farben analog zur SHERPA/RoMEO-Liste¹² angezeigt, die den Wissenschaftler auf einen Blick über die rechtlichen Standardbedingungen informieren.

Um die Eingabe von bibliographischen Daten für den Wissenschaftler auf ein absolutes Mindestmaß zu beschränken, sind bestimmte Import- und Exportmöglichkeiten unbedingt notwendig. Zuvorderst muss eine Import- und Exportfunktion für alle gängigen Literaturverwaltungsprogramme vorhanden sein, da der Wissenschaftler hier selbst bestimmen soll, ob er die Einträge in seinem Literaturverwaltungsprogramm oder im institutionellen Repository durchführen will. Bei der Vernetzung mit fachlichen Repositorien ist auf die jeweilige Arbeitsweise eines Wissenschaftlers, bzw. einer akademischen Disziplin, einzugehen. So besteht beispielsweise für Wirtschaftswissenschaftler der Universität Regensburg die Möglichkeit, Diskussionspapiere auf dem Publikationsserver zu veröffentlichen. Damit diese aber in der gesamten fachspezifischen Community gesehen werden, importiert man diese automatisiert in RePEc¹³. Anders hingegen verhält es sich im Fach Physik: Ein Großteil der Physiker bringt Veröffentlichungen in arXiv¹⁴ ein. Hier soll es möglich sein, dass Einträge in arXiv einfach mittels der arXiv-ID in das institutionelle Repository eingepflegt werden. Weitere fachliche Repositorien sollen auf analogem Wege eingebunden werden. Bei der Software EPrints sind solche Import- und Exportmöglichkeiten bereits vorhanden, bzw. können leicht technisch realisiert werden. Oftmals werden solche Zusatzprogramme von der internationalen Eprints Community bereits bereitgestellt. Beispielsweise wird eine Softwarelösung für den Import aus arXiv als Download angeboten. Herr Schallehn von der Universitätsbibliothek München hat uns ein File im ReDIF¹⁵-Format für den Export von bibliographischen Daten und Volltexten an RePEc zur Verfügung gestellt, welches nur noch angepasst werden musste.

Der Export von Publikationslisten für die Einbindung auf Internetseiten war von großer Bedeutung. Hierbei ist der Aspekt der Vergabe von Identifikationsnummern entscheidend, da über diese auf einfache Weise Literaturlisten erstellt werden können. Ein

6 Siehe <http://www.dspace.org/>

7 Siehe <http://roar.eprints.org/>, Stand 27.07.2010

8 Siehe <http://www.openoanr.org/>, Stand 27.07.2010.

9 Siehe <http://shibboleth.internet2.edu/>.

10 Siehe beispielsweise <http://www.openldap.org/>.

11 AJAX steht für Asynchronous JavaScript and XML. Für eine Einführung siehe <http://www.w3schools.com/ajax/default.asp>

12 Die Liste kann unter <http://www.sherpa.ac.uk/romeo/> eingesehen werden. Eine deutsche Version ist zugänglich unter <http://open-access.net/de/allgemeines/rechtsfragen/sherparomeolist/>

13 Siehe <http://repec.org/>

14 Siehe <http://arxiv.org/>

15 Siehe <http://ideas.repec.org/p/rpc/rdfdoc/redif.html>

weiterer wichtiger Punkt ist, dass die Listen in einem möglichst variablen Zitierformat auf den verschiedenen Seiten dargestellt werden können, um so den Anforderungen unterschiedlicher Wissenschaftsdisziplinen gerecht zu werden. Dazu ist ein Export in einem allgemeinen xml-Format notwendig. Die Einbindung in die unterschiedlichen Seiten erfolgt dann mittels eines php-Skripts, welches unter Benutzung von xsl den xml-Export in das gewünschte Zitierformat überführt. Eine Reihe von Zitierformaten kann auf den Seiten des Publikations-

bankens (wie z. B. PubMed¹⁶ oder ISI Web of Science¹⁷), da hier die importierten Daten auch mit Metadaten, wie Klassifikation und Keywords, angereichert sind.

Für die Veröffentlichung bestimmter Reihen war ein besonderer Wunsch, dass das Titelblatt automatisch generiert wird. So wird nach einem bestimmten Layout und aus ausgewählten Metadaten ein Titelblatt generiert und dieses als erste Seite in dem entsprechenden pdf-File der jeweiligen Einzelpublikation gespeichert.

Wie oben bereits angedeutet ist es mög-

also eine hohe Anpassungsfähigkeit haben. Diese Anpassungsfähigkeit muss bereits bei der Integration in ein gemeinsames System vorausgesetzt werden. Die gesamten Metadaten des bisherigen Hochschulschriftenservers müssen und können vollständig in EPrints abgedeckt werden.

Für diesen Zweck musste die standardisierte OAI-Schnittstelle erweitert werden. Zum Einen benötigen wir für die Vergabe von URNs die EPICUR¹⁸, zum Anderen für die Meldung der universitätseigenen Dissertationen XMetaDiss¹⁹. Diese konnten einfach in das System EPrints integriert werden. Dabei dienten uns die jeweiligen Programme von der Universitätsbibliothek München als Vorlagen, welche auf den Seiten²⁰ von EPrints heruntergeladen werden können.

Von Seiten der Bibliothek ist es noch wünschenswert, dass Redakteure nach ihren Aufgabengebieten verschiedene Rechte zugewiesen bekommen. Dies ist bereits in EPrints verwirklicht und die Rechte können individuell, detailliert und auf einfache Weise zugeteilt werden.

Zu erwähnen bleibt noch, dass mit der Ergänzungssoftware IRStats²¹ ein Statistikmodul für die Downloads von Volltexten zur Verfügung gestellt wird. Mit diesem Modul können die unterschiedlichsten Statistiken erzeugt werden und auch in verschiedene Internetseiten eingebunden werden.

Mit der Entscheidung für Eprints hat sich die Universitätsbibliothek Regensburg für eine international anerkannte Softwarelösung entschieden. Durch die große Anzahl der Institutionen, welche EPrints einsetzen und für ihre Zwecke anpassen, gibt es eine gute Möglichkeit, Erfahrungen²² sowie Programmteile²³ auszutauschen.

The screenshot shows the 'Universitätsbibliothek - Publikationen der Universität Regensburg' interface. The main content area displays the title 'Structural Conditional Correlation' by 'Weber, Enzo (2009)'. Below the title, there is a PDF viewer with a 'PDF - benötigt einen PDF-Betrachter wie iView, Ajaq oder Adobe Acrobat' message. The metadata section includes 'Institutionen: Wirtschaftswissenschaften - Institut für Volkswirtschaftslehre und Ökonometrie', 'Identifikationsnummer: 10001', 'Klassifikation: C32 Journal of Economics Literature Classification, G10 Journal of Economics Literature Classification', and 'Dewey-Dezimal-Klassifikation: 300 Sozialwissenschaften - 330 Wirtschaft'. The page also features a sidebar with subject categories and a right sidebar with options to export bibliographic data and create user profiles.

Abbildung 2: Übersichtsseite eines Eintrages beim Publikationsserver der Universität Regensburg.

onsservers heruntergeladen werden. Sollte ein Wissenschaftler ein bisher noch nicht vorhandenes Format benötigen, unterstützt die Universitätsbibliothek ihn selbstverständlich dabei.

Zu beachten ist dabei, dass durch die beschriebenen eindeutigen Identifikationsnummern eine Möglichkeit für das Durchführen von Suchen gegeben ist. Bereits bestehende Identifikationsnummern einzubeziehen ist zudem sehr sinnvoll, da auf diese Weise standardisierte Schnittstellen für die gesamte Universität geschaffen werden. So werden beispielsweise auf den Seiten der Wirtschaftswissenschaftlichen Fakultät Literaturlisten von Personen, Lehrstühlen etc. automatisch eingebunden. Dies erfolgt über die jeweiligen ID-Nummern. Besonders nützlich ist auch eine Importfunktion für bestimmte Inhalte von Daten-

banken, Embargoperioden zu definieren. Bei Artikeln, welche in Zeitschriften erschienen sind und die eine bestimmte Embargofrist haben, kann der Wissenschaftler das Ende dieser Frist angeben und der Volltext wird erst zu diesem Zeitpunkt vom System automatisch freigeschaltet. Somit muss der Kunde kein weiteres Mal das System benutzen, um den Volltext später hinzuzufügen.

Ein letzter wichtiger Entscheidungsgrund für das System EPrints ist dessen große Flexibilität. Als Dienstleister strebt die Universitätsbibliothek den optimalen Service für die Wissenschaftler der Universität an. Sollte es für Wissenschaftler einer bestimmten Fachrichtung spezielle Wünsche geben, so muss deren Realisierung maßgeblich zur Kundenzufriedenheit und Akzeptanz des Publikationsservers beitragen. Die Software muss

¹⁶ Siehe <http://ideas.repec.org/p/rpc/rdfdoc/redif.html>

¹⁷ Siehe <http://apps.isiknowledge.com/>

Integration des Hochschulschriftenservers in das institutionelle Repositorium

Bevor der Hochschulschriftenserver in das bestehende institutionelle Repositorium eingebettet werden konnte, mussten die Metadaten der dort eingestellten Publikationen erweitert werden. Die Datenstrukturen wurden analysiert und im institutionellen Repositorium um die fehlenden Elemente erweitert. Nachdem die Metadaten angepasst wurden, konnte zum Datentransfer ein Programm erstellt werden, welches die Daten-

¹⁸ Siehe <http://www.d-nb.de/wir/projekte/epicur.htm>

¹⁹ Siehe <http://www.d-nb.de/standards/xmetadiss/xmetadiss.htm>

²⁰ Siehe <http://files.eprints.org/119/>

²¹ Siehe <http://files.eprints.org/478/>

²² Zum Beispiel besteht eine E-Mail-Liste unter <http://files.eprints.org/view/type/release.html>

²³ Zusätzliche Programme können unter <http://files.eprints.org/> eingestellt und dort heruntergeladen werden.

bank des Hochschulschriftenservers ausliest und in ein für EPrints importfähiges xml-File schrieb. Dieses wurde dann in EPrints eingelesen. Wichtig dabei war, dass alle Daten aus dem Hochschulschriftenserver im neuen System gespeichert werden.

Schließlich mussten noch Hilfeseiten zum Onlinepublizieren, dem neuen System entsprechend angepasst und eingepflegt werden.

Einstellen des alten Services

Um einen reibungslosen Übergang des Hochschulschriftenservers in den Publikationsserver zu gewährleisten, wurde im ersten Schritt die Möglichkeit zum Veröffentlichen abgeschaltet. In einem zweiten Schritt wurden Weiterleitungen von allen Einträgen im Hochschulschriftenserver mittels der URN zum Publikationsserver eingerichtet. Somit blieb gewährleistet, dass alle Zitate und Links auf die alten Seiten weiterhin erreichbar sind.

Im Rahmen der Open Access Woche 2009 wurde der neue Publikationsserver vom Rektor der Universität Regensburg, Herrn Prof. Dr. Thomas Strothotte, offiziell eingeweiht.

Zusammenfassung

Mit der Wahl der Software EPrints hat sich die Universitätsbibliothek entschlossen, ein international bewährtes System einzusetzen.

Durch die Zusammenlegung des Hochschulschriftenservers mit den bisherigen institutionellen Repositorien wurde eine benutzerfreundliche Veröffentlichungsplattform geschaffen.

Dank

Wir bedanken uns bei der Universitätsbibliothek München, Herrn Volker Schallehn, für die Bereitstellung der einzelnen Programme.

AUTOREN

DR. RAFAEL BALL

Direktor der Universitätsbibliothek Regensburg
 Universitätsbibliothek
 Universitätsstraße 31
 93053 Regensburg
 rafael.ball@bibliothek.uni-regensburg.de



DR. GERNOT DEINZER

Fachreferent für
 Mathematik, Physik und
 Informatik
 Open Access Beauftragter
 der Universität Regensburg
 Universitätsbibliothek
 Regensburg
 93042 Regensburg
 gernot.deinzer@bibliothek.uni-regensburg.de



STAR-Kundentreffen 2010
Münchner Künstlerhaus
vom 13. bis 14. Oktober 2010

Integrierte Knowledge Center Lösungen

Bibliotheken, Archive, Dokumentations- und Informationszentralen, Museen und Landtage werden mit Anforderungen konfrontiert, die sich schnell verändern und stetig wachsen. Die **Cuadra STAR** Information Management Suite hat sich in 25 Jahren immer neu definiert, um dieser rasanten Entwicklung von Technologie und Anwendererwartung stets gerecht werden zu können.

- ▶ Archivmanagement
- ▶ Bibliotheksverwaltung
- ▶ Bild- und Medienarchiv
- ▶ Dokumentenmanagement
- ▶ eGovernment
- ▶ Integrierter Document Delivery Service
- ▶ Literaturverwaltung
- ▶ Museumsmanagement
- ▶ Normenverwaltung
- ▶ Parlamentsdokumentation
- ▶ Patentinformationsverwaltung
- ▶ Thesaurusmanagement
- ▶ Wissensmanagement
- ▶ Zeitschriftenverwaltung

Wir bieten Ihnen für Ihre individuellen Anforderungsprofile übersichtliche und anwenderfreundliche Lösungen!

GLOMAS Deutschland GmbH
 Germaniastraße 42
 80805 München

Fax 089 36 11 066
 Tel. 089 3 68 19 90

sales@glomas.de
www.glomas.com