

Open-Source-Software zur Realisierung von Institutionellen Repositorien – Überblick

Institutionelle Repositorien sind ein wesentliches Instrument, um Open Access sowohl auf dem grünen als auch auf dem goldenen Weg in die Praxis umzusetzen. Der vorliegende Artikel gibt einen Überblick über die derzeit in Deutschland eingesetzten Softwaresysteme, mit denen Institutionelle Repositorien technisch umgesetzt werden können und beleuchtet diese Lösungen unter den Anforderungen des DINI-Zertifikats für Dokumenten- und Publikationsservices 2007. Speziell werden erläutert: Opus, DSpace, EPrints, MyCoRe, Fedora.

Institutional repositories are a fundamental instrument for the implementation of open access using either the green road or the golden road. This article provides an overview of the software systems currently being used in Germany to operate institutional repositories and sheds light on these systems in terms of the currently valid (2007) certification for document and publication services established and administered by the non-profit association DINI (German Initiative for Network Information). Software given particular attention includes Opus, DSpace, EPrints, MyCoRe, and Fedora.

EINFÜHRUNG

Die Open-Access-Idee verlangt nach einer praktischen Realisierung. Neben dem goldenen Weg über verlagsbetreute Open-Access-Zeitschriften sind Institutionelle Repositorien das Instrument, dessen sich vorrangig wissenschaftliche Einrichtungen wie etwa Universitäten bedienen, um ihre Publikationen frei zugänglich bereitzustellen. Dies soll jedoch keineswegs bedeuten, dass nicht auch die Bereitstellung von reinen verlagsbetreuten Open-Access-Zeitschriften über ein Institutionelles Repository realisiert werden kann.

Der vorliegende Artikel gibt einen Überblick über die derzeit in Deutschland eingesetzten Systeme, stellt deren Besonderheiten heraus, gibt eine Übersicht über deren Leistungsspektrum und beleuchtet den Aufwand bei der Einführung und Wartung. Dabei ist es wichtig, die Vor- und Nachteile für unterschiedliche Anwendungsansprüche herauszustellen. Zudem sollen die Systeme unter den technischen Aspekten des DINI-Zertifikats für Dokumenten- und Publikationsservices 2007 betrachtet und miteinander verglichen werden.

Auffällig ist, dass es bei weitem nicht nur die Universitätsbibliotheken oder Rechenzentren sind, die Institutionelle Repositorien aufsetzen, sondern dass eine ganze Reihe von Fachbereichen, Fachgesellschaften und Forschungseinrichtungen eigene Institutionelle Repositorien aufbauen. Daher sind auch die Ansprüche, die aus organisatorischer und technischer Sicht an die Software gestellt werden, sehr unterschiedlich und somit auch die Wahl des konkreten Systems sehr

stark davon abhängig, inwieweit einzelne Aspekte für die Betreibereinrichtung Priorität besitzen. Zudem haben sich speziell in Deutschland Lösungen etabliert, die zielgerichtet auf die Belange deutscher Universitätsbibliotheken zugeschnitten sind und somit im internationalen Vergleich wenig Berücksichtigung finden.

ASPEKTE ZUR BEWERTUNG EINZELNER SYSTEME ZUR REALISIERUNG INSTITUTIONELLER REPOSITORIEN

Im Projekt »Open Access Repositories in New Zealand« (OARINZ) [1] wurden 2006 einige Open-Source-Systeme zur Realisierung Institutioneller Repositorien aus dem OSI Guide [2] unter verschiedenen Gesichtspunkten bewertet. Zu den in die endgültige Auswahl gehörenden Systemen haben es allerdings nur Fedora, EPrints und DSpace geschafft, die eine entsprechende weltweite Nutzergemeinschaft haben und deren gesamter Sourcecode auch als Open Source verfügbar ist. Betrachtet man die im deutschen Raum eingesetzten Systeme, so ist die obige Auswahl nicht repräsentativ für unseren nationalen Kontext. Insofern werden in diesem Artikel weitere Systeme mit größerer Verbreitung in Deutschland in die Betrachtung aufgenommen. Diese sind: OPUS und MyCoRe. Da derzeit nur zwei Fedora-Implementierungen in Deutschland betrieben werden, wird das System nicht an vorderer Stelle betrachtet.

Zu den allgemeinen Aspekten, die zur Bewertung in OARINZ herangezogen werden, gehören z. B. die Einfachheit der Installation und Wartung des Systems sowie dessen Fähigkeit, mehrere Installationen gleichzeitig auf einer Hardware/Software-Plattform laufen zu lassen, um damit ggf. Hosting Services realisieren zu können. Des Weiteren spielt die Interoperabilität mit anderen Systemen und das Vorhandensein entsprechender Schnittstellen eine große Rolle, um z. B. das Institutionelle Repository in die Informationsinfrastruktur der Einrichtung zu integrieren, sei es in die einer Universität, einer Verbundzentrale oder einer wissenschaftlichen Institution. Dementsprechend kann es lokal von Bedeutung sein, das Institutionelle Repository in den OPAC zu integrieren, in Lehr- und Lernumgebungen oder in das lokale Nutzermanagement (LDAP) einzubinden bzw. an Systeme zur Leistungsmessung anzubinden. Weitere Aspekte sind die



Foto: privat

Susanne Dobratz

größte Verbreitung in Deutschland: OPUS und MyCoRe

Vielfalt von Institutionellen Repositorien

Skalierbarkeit des Systems, die Sicherheit des Systems, die Stärke der dahinterstehenden Nutzergemeinschaft (Community), die Art der Open-Source-Lizenz und der Schwierigkeitsgrad, mit dem eigene Erweiterungen der Code-Basis realisiert werden können, um Anpassungen an lokale Gegebenheiten vorzunehmen.

Die Beantwortung der Frage jedoch, inwieweit sich ein einzelnes Softwarepaket eignet, um lokal ein Institutionelles Repository aufzubauen, hängt immer von der lokalen Zielrichtung und den eigenen Prioritäten ab. Im Hinblick auf die bereits entwickelte und weiter auszubauende Infrastruktur von Institutionellen Repositorien in Deutschland, wie sie im DFG-Projekt OA-Netzwerk [5] erfolgen soll, stellen die technischen Anforderungen des DINI-Zertifikats 2007 [4] einen wichtigen Ansatzpunkt dar. Die folgenden technischen Anforderungen wurden daraus extrahiert und fließen in den darauf folgenden Vergleich der Softwaresysteme ein.

ÜBERBLICK ÜBER SYSTEME

ROAR¹, das von der Universität Southampton betreute Registry of Open Access Repositories, listet derzeit

79 deutsche Dokumentenserver auf. Die DINI-Datenbank enthält derzeit 109. Von den aufgeführten deutschen Institutionellen Repositorien haben 19 das DINI-Zertifikat 2004 errungen, das Zertifikat 2007 wurde noch nicht vergeben, es befinden sich aktuell 3 Anträge in der Bearbeitung.

Es ist kaum möglich, einen genauen Überblick über die derzeit existierenden Institutionellen Repositorien in Deutschland zu erhalten, entstehen doch ständig neue, die nicht notwendigerweise an eine Universitätsbibliothek oder eine ähnliche Einrichtung gekoppelt sind. Oft betreiben Einrichtungen sogar mehrere Server, die dann unterschiedlichen Publikationsstrategien dienen: Medienserver zur Verwaltung multimedialer Lehr- und Lernmaterialien, Dissertationsserver, Repositorien, die nur der Aufnahme von Pre- oder Postprints wissenschaftlicher Arbeiten dienen, Repositorien für verschiedene Nutzergruppen, etwa Graduiertenkollegs, Institute etc. Belastbare exakte Zahlen sind auch deshalb schwer zu ermitteln, weil die Institutionellen Repositorien oftmals Teil eines größeren Angebots z. B. eines Fachportals sind und nicht als separates Institutionelles Repository auftreten oder

Upload – [4] Punkt 2.3: Bereitstellung von Online-Formularen für Autoren, die ihre Pre- und Postprints einbringen wollen (Document upload tools)

Rechte – [4] Punkt 2.4: Eine Lizenzierung des Werkes durch den Autor über entsprechende Vereinbarungen (CC, DPPL) wurde oder wird vorgenommen.

Dokumentation – [4] Punkt 2.5.1: Es existiert eine Dokumentation des technischen Systems mit relevanten Versionsangaben und technischen Parametern zu allen Komponenten.

Nachweis – [4] Punkt 2.5.1 Kontrollierte und nachweisbare Aufnahme von Dokumenten aus technischer Sicht.

PI – [4] Punkt 2.5.2: Verwendung von Persistent Identifiers, dazu zählen Systeme, die einen Resolver-Dienst besitzen, z. B. urn:nbn oder DOI.

Keywords – [4] Punkt 2.6.1: Verbale Sacherschließung durch freie Schlagwörter oder klassifikatorische Erschließung wird durchgeführt (technisch unterstützt).

DDC – [4] Punkt 2.6.1: Dewey-Dezimalklassifikation (DDC) gemäß der Verwendung in der Deutschen Nationalbibliografie als allgemeine klassifikatorische Erschließung aller Dokumente (entsprechend den DINI-OAI-Empfehlungen) wird angewandt.

Metadaten – [4] Punkt 2.6.2: Metadaten werden frei zugänglich angeboten. Metadaten sind nach Dublin Core Simple (ISO 15836:2003) strukturiert.

OAI – [4] Punkt 2.6.2: OAI-PMH 2.0 entsprechend den DINI-OAI-Empfehlungen wird unterstützt.

WWW – [4] Punkt 2.6.3: Webserverschnittstelle für Nutzer ist vorhanden.

Statistik – [4] Punkt 2.7: Jeder Dokumentenserver muss im Rahmen der rechtlichen Bestimmungen eine eigene konsistente Zugriffsstatistik führen (Webserver-Log).

LZA – [4] Punkt 2.8: Dauerhafte Verbindung der Metadaten mit den Dokumenten (z. B. Verbindung über Persistent Identifier oder zusammen in einem Container).

Aspekte, die von den Systemen a priori nicht umgesetzt werden, müssen lokal mit anderen Software-Lösungen realisiert werden. Ein Beispiel ist das Führen einer dokumentierten Zugriffsstatistik, die ebenso über gängige Systeme zur Webstatistik wie *Webalizer* angeboten werden kann, ohne dass sich dies negativ auf die Vergabe des DINI-Zertifikats auswirkt.

sich nicht bei den entsprechenden Verzeichnissen angemeldet haben.

OPUS

Das Online-Publikationssystem der Universität Stuttgart (OPUS) wurde 1998 von der Universitätsbibliothek und dem Rechenzentrum der Universität Stuttgart mit einer Förderung durch den DFN-Verein entwickelt. Seitdem wird die Weiterentwicklung und die Herausgabe neuer Versionen durch die Universität Stuttgart und das Bibliotheksservicezentrum Baden-Württemberg (BSZ) betrieben. Bei Nutzung von OPUS für den produktiven Betrieb eines eigenen institutionellen Repositoriums fallen derzeit einmalig 250 EUR Nutzungsgebühren an. Ursprünglich für das Management elektronischer Publikationen an der Universität Stuttgart entwickelt, existieren heute mehr als 60 Installationen in Hochschulen und Verbundzentralen, wobei besonders hervorzuheben ist, dass das Bibliotheksservicezentrum, der Bayerische Bibliotheksverbund, das Hochschulbibliothekszentrum Köln sowie der Kooperative Bibliotheksverbund Berlin-Brandenburg ein Hosting für Institutionelle Repositorien auf der Basis von OPUS anbieten.

Da es speziell für den Einsatz an deutschen Hochschulen konzipiert wurde, unterstützt es die wichtigsten Schnittstellen und Standards wie die Generierung von Uniform Resource Names (URN) in Form der National Bibliographic Number (NBN) entsprechend den Richtlinien der Deutschen Nationalbibliothek (DNB) über eine XEPICUR-Schnittstelle, sowie XMetaDISS als Format zur Meldung von Online-Hochschulschriften an die DNB. Zudem stellt es standardmäßig eine Proprint-Schnittstelle [6] für einen Print-On-Demand-Dienst zur Verfügung. Last but not least stellt OPUS für die Autoren ein Modul zur Definition von Nutzungslizenzen, u. a. Creative-Commons-Lizenzen bereit.

OPUS zeichnet sich vor allem dadurch aus, dass die Installation und Wartung des Programmpaketes auf einem Standard-LINUX-System relativ einfach und schnell ist. Dabei unterstützt OPUS vor allem den Publikationsprozess in der Bibliothek: der Autor lädt sein Dokument in das System, dort überprüft ein Administrator die Richtigkeit und Vollständigkeit, korrigiert ggf. die Metadaten und schaltet das Dokument für die Open-Access-Bereitstellung über das WWW-Interface, dessen Ausgabesprache konfigurierbar ist, frei. Dabei werden die Vergabe von URNs, die Meldung an die DNB sowie eine Schlagwortvergabe unterstützt. Es werden verschiedene Browsingmöglichkeiten sowie die Suche in den Metadaten und den Volltexten (über Google) angeboten. OPUS erlaubt die Definition differenzierter Zugriffsbeschränkungen für Dokumente

(Welt/Campus/weitere). Es ermöglicht des Weiteren die Verwaltung von Schriftenreihen mit Bandzählung und von Sammlungen mit beliebig gestuften Hierarchien, allerdings unter einem einheitlichen Layout.

OPUS ist mit neun DINI-zertifizierten Dokumenten-Servern und insgesamt ca. 60 Systemen die derzeit meistgenutzte Software zur Realisierung eines Institutionellen Repositoriums in Deutschland. Das OPUS-Programmpaket ist mittels PHP und MySQL realisiert, was zwar vom programmiertechnischen Stand aus nicht unbedingt die neueste Technologie im Sinne des »Web2.0« repräsentiert, aber erprobt, stabil und skalierbar ist.

Aktuelle Version: 3.1

Systemplattform: GNU LINUX und andere UNIX Derivate, Apache, PHP, MySQL

Lizenz: Universität Stuttgart (250 EUR einmalige Nutzungsgebühr)

Support: Wiki, Bugtracker, siehe http://elib.uni-stuttgart.de/opus/doku/opus_sw.php

Ca. 60 **Installationen** in Deutschland, davon 9, die das DINI-Zertifikat 2004 tragen, siehe [8].

DSpace

Das Massachusetts Institute of Technology (MIT) entwickelte DSpace in Kooperation mit der Firma Hewlett-Packard zwischen März und November 2002. Die erste verfügbare Version (1.2) wurde im April 2004 bereitgestellt. Das System läuft bei MIT im Produktionsbetrieb, mit der DSpace-Föderation haben sich weitere große Forschungseinrichtungen zusammengeschlossen, um die technologische Weiterentwicklung des Systems zu betreiben. Neben der Föderation stellt das MIT den Sourcecode des Systems für Weiterentwicklungen zur Verfügung und organisiert die Herausgabe neuer Versionen, bei denen einzelne Erweiterungen dann in die Standard-Distribution aufgenommen werden. Interessant ist, dass die Universität Kassel das System so erweitert hat, dass es neben dem im angelsächsischen Raum primär eingesetzten Handle-System auch das URN-Format der DNB unterstützt.

DSpace unterstützt die gängigen Schnittstellen: OAI, handle.net, URN, OpenURL. Es stellt ferner Schnittstellen bereit, um Creative-Commons-Lizenzen vergeben zu können, und erlaubt standardmäßig ein Browsing über die Sammlung sowie eine Suche in Metadaten. Bei der Installation, Anpassung und Wartung des Systems benötigt man einen Java-Programmierer in seinem Personalkontingent. Auch ist die Erstinstallation ein wenig komplexer als zum Beispiel bei

OPUS derzeit meistgenutzte Software

Sourcecode steht zur Weiterentwicklung zur Verfügung

ausgereiftes
Nutzermanagement

OPUS, aus technologischer Sicht ist das System zukunftsträchtiger durch den Einsatz von Java-Technologie, die es erlaubt, bei der Programmierung und Erweiterung des Systems strukturierter vorzugehen und so Anpassungen an neue aktuelle Webtechnologien sauberer und strukturierter vorzunehmen.

Das System wurde entworfen, um den intellektuellen Output einer multidisziplinären Organisation zu erfassen und zu verwalten. Insofern stellt es ein ausgereiftes Nutzermanagement bereit, bei dem rollenbasiert der Publikationsworkflow vom Upload in das System bis hin zur Freischaltung (Campus, weltweit etc.) gesteuert wird.

Hervorzuheben ist, dass DSpace die Struktur der lokalen Anwendergemeinde nutzt, um eine entsprechende Systemstruktur zu generieren. Das heißt, einzelne Institute, Forschungszentren, Fakultäten, die Teil einer großen Forschungseinrichtung wie etwa einer Universität sind, erhalten einen eigenen Bereich für ihre Dokumente und für deren Verwaltung im System. Damit verbunden sind auch gruppenspezifische Anpassungen bezüglich der Autorisierung, der Urheber- und Verwertungsrechte und des Zugriffs. Bei der Planung des Systems bestand ursprünglich der Anspruch, eine echte Langzeitarchivierung im Sinne des Open Archival Information System (OAIS) [7] zu realisieren, in der aktuellen Standarddistribution ist dies jedoch noch nicht umgesetzt.

DSpace ist skalierbar, gut nutzbar für größere Forschungseinrichtungen, die eine Vielfalt von Publikationsstrategien verfolgen und natürlich einsetzbar, um den grünen Open-Access-Weg der Pre- und Postprints zu unterstützen. Es eignet sich auch als Hosting-System, allerdings gibt es in Deutschland derzeit noch keine Institutionen, die dies anbieten. International sind über 100 Installationen verzeichnet.

EPrints: für Eigenpublikationen von Wissenschaftlern optimiert

Aktuelle Version: 1.4.2 (Mai 2007)

Systemplattform: UNIX/LINUX, auch Windows (2000,NT,XP), OS X, Apache Webserver, Tomcat, Java Server Pages-Technologie (JSP), Programmcode in Java, PostgreSQL-Datenbank (alternativ Oracle), Lucene Suchmaschine (Apache Project), Apache Webserver, Tomcat als Java Servlet Engine, Jena (RDF Toolkit von HP Labs), OAICat von OCLC

Lizenz: Open Source (BSD)

Support: DSpace Federation, Bug-Tracker, WIKI,FAQ, User Group Meetings, Sourcecode verfügbar über Sourceforge.net

Von den aufgeführten 6 **Installationen** in Deutschland besitzt KOBRA das DINI-Zertifikat 2004:

KOBRA (Kasseler OnlineBibliothek, Repository und Archiv) (<https://kobra.bibliothek.uni-kassel.de>)

DSpace at ICSY/ Integrated Communication Systems Lab der TU Kaiserslautern FB Informatik (<http://DSpace.icsy.de>)

ELDORADO (UB Uni Dortmund) (TU Kaiserslautern)

Geo-LEO/ Publikationsplattform für geo- und montanwissenschaftliche Informationsressourcen (<http://geoleoedocs.sub.uni-goettingen.de:8080/DSpace>)

Institute for Political Science, University of Duisburg-Essen (<https://pol-serv1.uni-duisburg.de/DSpace>)

Forschungszentrum Jülich (<http://juwel.fz-juelich.de:8080/DSpace122fzj>)

EPrints

EPrints ist eine Entwicklung der University of Southampton aus dem Jahr 2000 und wird auch von dort gewartet und versioniert. Es existieren über 200 Installationen weltweit. Die Software wird derzeit auf zwei der DINI-zertifizierten Institutionellen Repositorien eingesetzt, an der Ludwig-Maximilians-Universität München (LMU) sogar für 4 Installationen: edoc – für digitale Dissertationen, epub – für elektronische Publikationen der LMU, ebook – für Ebooks sowie MPRA – das RePEc Archiv der LMU. Im Gegensatz zu den anderen Systemen ist EPrints für die Eigenpublikationen von Wissenschaftlern (Self Archiving) optimiert, wird jedoch auch zum Betrieb von Institutionellen Repositorien eingesetzt. EPrints unterstützt den Publikationsprozess für Autoren, Herausgeber und Institutionen. Es verfügt daher über ein Rechtemanagement. EPrints stellt eine OAI-Schnittstelle bereit sowie standardmäßig einen Export von Metadaten in den Formaten MODS, METS, diverse Referenzmanager, BiBTeX, OpenURLContextObject. Es unterstützt URNs. Eine Alerting-Funktion informiert über die Bereitstellung neuer Dokumente im Repositorium, die History-Funktion ermöglicht Administratoren das Verfolgen von Änderungen.

Diese Software setzt Unix/Linux voraus und nutzt einen Apache http-Server. Um EPrints zu installieren, benötigt man keine Spezial- oder Programmierkenntnisse, um es erweitern oder anzupassen benötigt man Programmierkenntnisse in Perl und MySQL. Grundsätzlich sind mehrere Archive unter einer Installation lauffähig, d. h. ein Hosting ist grundsätzlich möglich. Es gibt derzeit jedoch keine Einrichtung in Deutschland, die ein Hosting auf EPrints-Basis anbietet. Die Standarddistribution ist für den englischsprachigen Bereich optimiert und muss daher bei der Installation an die deutschen Nutzer angepasst werden. Das be-

inhalte auch die Konfiguration der OAI-Schnittstelle für den Export von Metadaten im XEPICUR- oder XMetaDISS-Datenformat.

Aktuelle Version: 3.0 (Januar 2007)

Lizenz: Open-Source (GPL)

Systemplattform: Linux / Solaris, Apache, MySQL, Perl (mod_perl)

Support: Wiki: (<http://wiki.EPrints.org>), Datei-Archiv (<http://files.EPrints.org>), Mailingliste: (<http://www.EPrints.org/tech.php>)

11 Installationen in Deutschland:

DLR electronic library (<http://elib.dlr.de>)

Universitäts- und Landesbibliothek Darmstadt (<http://ebooks.ulb.tu-darmstadt.de>)

gdea – Graph Drawing E-print Archive (GDEA) Universität Köln, FB Informatik (<http://gdea.informatik.uni-koeln.de>)

LMU München, Bibliothek betreibt 4 Server:

— Dissertationen (<http://edoc.uni-muenchen.de>),

— E-Publikationen (<http://epub.uni-muenchen.de>),

— E-Books (<http://ebooks.uni-muenchen.de>)

— LMU München, Munich Personal RePEc Archive (MPRA) (<http://mpra.uni-muenchen.de>)

TU Berlin, Zentrum für Multimedia in Lehre und Forschung (www.mulf.tu-berlin.de)

UB Uni Regensburg (<http://epub.uni-regensburg.de>)

Universität Oldenburg, Graduiertenkolleg Trustsoft (<http://trustsoft.uni-oldenburg.de:19080>)

DLR Clearingstelle für Verkehrsdaten und Verkehrsmodelle (<http://modelle.clearingstelle-verkehr.de>)

MyCoRe

MyCoRe wurde zunächst im Projekt Miles an der Universität Essen entwickelt. Ursprünglich stark an die IBM Hardware und Software »IBM Digital Library« (derzeitige Variante wird als *IBM Content Manager* vertrieben) gebunden, hat sich das Produkt MyCoRe zu einem Open-Source-Produkt entwickelt, das in einem Konsortium mehrerer Universitäten weiter entwickelt und gepflegt wird. MyCoRe ist eher ein Rahmen, bestehend aus verschiedenen Software-Paketen und einem speziellen Management-System, das es erlaubt, digitale Bibliotheken, seien es Institutionelle Repositorien, Lehr- und Lernserver, Medienserver etc. zu managen. Für alle Grundfunktionen einer digitalen Bibliothek will MyCoRe Implementierungen bereitstellen, so dass einfache Anwendungen primär durch Konfiguration statt Programmierung realisiert werden können. Dabei sind die einzelnen Pakete hochgradig konfigurier-

bar und anpassbar. Es wird eine Basisvariante bereitgestellt, die so gut wie keine Programmierkenntnisse verlangt und in der Nutzer bereits eigene Konfigurationen vornehmen können. Besonders hervorzuheben ist die Möglichkeit, das Datenmodell vollständig selbst zu konfigurieren. So lassen sich Sammlungen, Zeitschriften, Bände etc. abbilden.

Die Kernsoftware verfügt über alle wichtigen Funktionen, die ein Institutionelles Repositorium anbieten sollte: Unterstützung des Publikationsprozesses vom Autorenpupload zur Klassifikation, Konfigurierbarkeit der Oberfläche und des Datenmodells über XML, Bereitstellung wichtiger Schnittstellen wie URNs, OAI, XMetaDiss, XEPICUR, Z39.50, MAB2, SOAP. Es unterstützt ein Video-Streaming sowie eine Suche über Metadaten und Volltexte, selbst bei verteilten MyCoRe-Anwendungen. Damit ist das System auf den deutschen Raum spezifiziert und kann ohne größere Anpassungen sofort eingesetzt werden. Die Voraussetzung für einen längerfristigen Einsatz an einer Einrichtung, will man das Potenzial des Systems vollständig ausnutzen, ist allerdings die Bereitstellung von Programmierkapazitäten für Java-basierte Technologien.

MyCoRe ist ein hochkomplexes und modulares System, es eignet sich zur Bereitstellung mehrerer Installationen über ein System und kann somit auch ein Hosting realisieren.

Die technische Basis des Systems bilden Java-Klassenbibliotheken, XML-Techniken und unterschiedliche Datenbank-Backends wie IBM Content Manager und IBM DB2, sowie MySQL, hsql, Oracle und XMLDB-Datenbanken. Dies ermöglicht es, wirklich große und skalierbare Anwendungen zu betreiben.

Aktuelle Version: 1.3.0 (Januar 2007)

Systemplattform: Windows, Linux, AIX, Solaris, Mac OS, Java 5 SDK, Standard Edition ausreichend, Java Web Application Server (jetty:// integriert, auch tomcat, Websphere), Relationale Datenbank (hsql integriert, auch mysql, Oracle, DB2), Java- und XML/XSL Basis,

Lizenz: Open Source (GNU GPL),

Support: Download über mycore.de, Sourceforge, CVS Repository, Mailingliste, Bug-Tracking, Wiki, IRC Kanal

15 Installationen in Deutschland, siehe www.mycore.de unter »Anwendungen«:

www.mycore.de/content/main/anwendungen.xml?XSL.lastPage.SESSION=/content/main/anwendungen.xml

MyCoRe ist auf den deutschen Raum spezifiziert

**hervorragende Eigenschaft:
RDF-Fähigkeit**

Fedora

Fedora ist eine Entwicklung der University of Virginia und der Cornell University. Es stellt eher eine Architektur (Fedora: Flexible Extensible Digital Object and Repository Architecture) bereit als ein fertiges Komplettsystem, das Out-of-the-box installiert werden kann. Fedora ist sowohl im Quellcode als auch als binäre Distribution erhältlich. Beide Installationsarten sind gut beschrieben und einfach durchzuführen. Die Software ist in Java entwickelt und sticht insbesondere durch die RDF-Fähigkeit des Gesamtsystems hervor. Fedora unterstützt die Verwaltung von Objekten bzw. auch nur die Beschreibung von Objekten. Der Import und der Export von Objekten in das Repositorysystem kann entweder über METS (Metadata Encoding and Transmission Standards) oder über das Fedora eigene Format (Fedora Object XML: FOXML) realisiert werden. Zukünftig sollen die Digital Item Declaration Language MPEG2/DIDL und METS 1.3 unterstützt werden.

Fedora stellt drei SOAP-basierte Webservices bereit:

- eine Management API, die eine Administrationschnittstelle zur Administration des Repositoriums bereitstellt, die für Nutzer notwendig sind, um Objekte zu erstellen und zu warten.
- Eine Access API, die den Zugang zu den Objekten im Repository und die Verteilung der Objekte unterstützt.
- Eine abgespeckte Version des Access API, die als http-basierter Webservice funktioniert.

Dieses sind die einzigen von der Software bereitgestellten Nutzerschnittstellen und Möglichkeiten, um auf die Objekte und Funktionen zuzugreifen. Alles andere muss man selbst programmieren. Daher ist Fedora auch keine Software, die direkt nach der Installation eingesetzt werden kann. Sie setzt immer einen Aufwand für die Integration in bereits bestehende oder neu zu erstellende Anwendungen voraus.

In Deutschland wird Fedora beim Projekt Digital Peer Publishing (DIPP-NRW) (www.dipp.nrw.de) eingesetzt, über das ein Hosting-Service in Anspruch genommen werden kann. Die Max Planck Digital Library arbeitet daran, Fedora im eSciDoc-Projekt für die Max-Planck-Gesellschaft einsatzfähig zu machen.

Aktuelle Version: 1.2.1

Systemplattform: Apache, Tomcat 5, MySQL, Oracle 9i, OpenSource DB McKoi, Java

Lizenz: Mozilla Public License Version 1.1

WEITERE SYSTEME UND EIGENENTWICKLUNGEN

CDSware

Die CERN Document Server Software (CDSware) wurde speziell für das Institutionelle Repositorium des CERN entwickelt, ist inzwischen als Open Source frei verfügbar und in der Lage, sehr große Sammlungen von Dokumenten, bibliographischen Datensätzen und Daten zu verwalten. Es erlaubt eine kontrollierte Aufnahme von Dokumenten, unterstützt eine Qualitätskontrolle und unterschiedliche Veröffentlichungsstrategien. In Deutschland setzt das HBZ Köln die Software für MeIND (www.meind.de) ein.

ETD-DB

ist eine Entwicklung der Virginia Polytechnic Institute and State University (VT) zur Verwaltung der eigenen Sammlung digitaler Dissertationen. Sie stellt alle Schnittstellen bereit, die das DINI-Zertifikat 2004 fordert.

Der DuetT-Server der Universität Duisburg Essen: DuetT: Dokumentenserver Campus Duisburg (www.ub.uni-duisburg.de/ETD-db) nutzte diese Software, durch die Fusion mit dem miLess-System wird der Server jedoch nicht aktiv weiter entwickelt.

Auch Systeme und Angebote, wie Living Reviews, Open Repository oder Open Journal System, deren Fokus auf der Unterstützung des Publikations- und Begutachtungsprozesses von Open Access Journals liegt, finden zunehmend Anwender.

Living Reviews

Das Living Reviews ePublishing Toolkit wird im Rahmen des Living-Reviews-Projekts am Max-Planck-Institut für Gravitationsphysik entwickelt und ist ein Softwarepaket, mit dessen Hilfe Online Journals als sogenannte Webapplikation erstellt, gewartet und angeboten werden können. Beispiele sind die Zeitschriften Living Reviews in Relativity des Max-Planck-Instituts für Gravitationsphysik Potsdam (<http://relativity.livingreviews.org>) sowie Living Reviews in Solarphysics des Max-Planck-Instituts für Sonnensystemforschung (<http://solarphysics.livingreviews.org>).

Open Journal System

Das vom kanadischen Public Knowledge Project entwickelte System ist ein Open Source Online Journal Management- und Publikationssystem. Es stellt verschiedene Funktionen zum Aufbau eines Open Access Journals bereit und zum Management des Reviewing-Prozesses. Es stellt u. a. eine OAI2.o-Schnittstelle bereit. In

**Living-Reviews-Projekt am
Max-Planck-Institut für
Gravitationsphysik**

Deutschland wird es z. B. angewandt an der Universität Hannover, für das Journal of Academic Legal Studies (www.joals.org//index.php/current).

Einige DINI-zertifizierte Dienste nutzen Eigenentwicklungen, so etwa die Humboldt-Universität zu Berlin mit dem edoc-Server, die auch den edoc-Server der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) auf dem System hostet, ask23: das Archivsystem der HFBK Hamburg (<http://ask23.hfbk-hamburg.de/draft>), der Webdoc-Server der Georg-August-Universität Göttingen (<http://webdoc.sub.gwdg.de>), der Hochschulschriftenserver der Sächsischen Universitäts- und Landesbibliothek (<http://hsss.slub-dresden.de/oai/index.htm>) sowie die elib der Universität Bremen (<http://elib.suub.uni-bremen.de>).

FAZIT

Institutionelle Repositorien lassen sich mit unterschiedlichen Technologien realisieren. Die in diesem Artikel vorgestellten Systeme sind diejenigen, die derzeit bei deutschen Dokumentenservern im Produktionsbetrieb eingesetzt werden.

Zahlreiche technische Lösungen sind mehr oder weniger ausgehend von einer lokalen Entwicklungsphilosophie für den Eigenbedarf einzelner Dokumentenserver entstanden. In den oben genannten Fällen ist die Entwicklung lokaler Software zu einer Produktentwicklung gereift, die die Nachnutzung an anderen Standorten unter anderen Bedingungen gestattet und so zu einer erfolgreichen Nachnutzung bis hin zur Bildung einer Gemeinschaft zur Pflege und Weiterentwicklung der Software geführt hat.

DSpace unterstützt community-basierte Einreichungsrichtlinien und Prozesse und soll später auch die Langzeitarchivierung dieser Objekte realisieren. EPrints wurde als schnelle und einfache Lösung zur Unterstützung des Grünen Weges zu Open Access entworfen, Fedora bietet ein System mit vielen Eigenschaften zum Aufbau einer digitalen Bibliothek, das gut mit sehr großen Repositorien umgehen kann. Bei den deutschen Entwicklungen zeichnet sich MyCoRe durch seine Flexibilität, sehr gute Konfigurierbarkeit und seine Anwendbarkeit für ganz verschiedene digitale Bibliotheken und Repositorien aus, wogegen OPUS eine sofort nutzbare out-of-the-box Lösung bereitstellt, die speziell auf die deutschen Anforderungen zugeschnitten ist und eine breite Anwendergemeinschaft besitzt.

Das DINI-Zertifikat »Dokumenten- und Publikationsservice« definiert im Hinblick auf die Systemauswahl den technischen Mindeststandard der in Hochschulen verwendeten Softwarelösungen. Die oben

untersuchten Systeme genügen alle (zum Teil nach entsprechender Konfiguration) den gestellten technischen Anforderungen. Daher wird an dieser Stelle auf eine tabellarische Darstellung verzichtet.

Sollte man die Auswahl zwischen diesen Systemen treffen müssen, so sind es immer die lokalen Besonderheiten, die letztendlich den Ausschlag für die Entscheidung geben müssen, so die Verfügbarkeit entsprechend in den Programmiersprachen ausgebildeten Personals bzw. die grundsätzliche Haltung zu einem integrierten Informations- und Medienmanagement vor Ort bzw. die Strategie im Umgang mit dem Institutionellen Repository und seine organisatorische Verankerung.

LITERATUR

[1] **Wyles, Richard** (2006): Technical Evaluation of Research Repositories. Open Access Repositories in New Zealand. <https://eduforge.org/docman/view.php/131/1062/Repository%20Evaluation%20Document.pdf> [Stand 02.07.2007]

[2] **Open Society Institute**: OSI Guide to Institutional Repository Software, Version 3.0 Edition, August 2004. www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf [Stand 02.07.2007]

[3] **Beier, Gerhard and Velden, Theresa**: The eDoc-Server Project: building an institutional repository for the Max Planck Society, High Energy Physics Libraries Webzine, issue 9, March 2004. <http://library.cern.ch/HEPLW/9/papers/4/> [Stand 02.07.2007]

[4] **DINI Arbeitsgruppe Elektronisches Publizieren**: DINI-Zertifikat Dokumenten- und Publikationsservice 2007, DINI-Schriften 3, herausgegeben von der Deutschen Initiative für Netzwerkinformation, September 2006, <http://nbn-resolving.de/urn:nbn:de:kobv:11-10068508> [Stand 02.07.2007]

[5] **OA-Netzwerk**: siehe Artikel von Uwe Müller und Peter Schirnbacher dazu in diesem Heft.

[6] **Mittler, Elmar und Schulz, Matthias**: ProPrint world-wide print-on-demand services for study and research (2004). Library Hi Tech 22 [2], Seite 227–230, DOI: 10.1108/07378830410543548

[7] **CCSDS (Consultative Committee for Space Data Systems)** (2002): Reference Model for an Open Archival Information System (OAIS). Blue Book ISO 14721:2003 Issue 1. www.ccsds.org/docu/dscgi/ds.py/Get/File-143/650xob1.pdf [Stand 02.07.2007]

[8] **Liste der DINI-zertifizierten Server**: www.dini.de/no_cache/service/dini-zertifikat/zertifizierte-server/

[9] **Freudenberg, Stefan**: OPUS Funktion und Entwicklung. Vortrag auf dem Workshop Open Access und das DINI Zertifikat 2007. Februar 2007. Frank-

lokale Besonderheiten sind für Systemauswahl stets ausschlaggebend

alle genannten Systeme genügen dem technischen Mindeststandard des DINI-Zertifikats

furt/Main. Vgl. auch http://de.wikipedia.org/wiki/OPUS_%28Software%29

[10] **Quast, Andreas:** Einsatz der Software DSpace am Beispiel von GEO-Leo edocs. Vortrag auf dem Workshop Open Access und das DINI Zertifikat 2007. Februar 2007. Frankfurt/Main

[11] **Schallehn, Volker:** EPrints 3. Vortrag auf dem Workshop Open Access und das DINI Zertifikat 2007. Februar 2007. Frankfurt/Main

[12] **Lützenkirchen, Frank:** Dokumenten- und Publikationsserver mit MyCoRe. Vortrag auf dem Workshop Open Access und das DINI Zertifikat 2007. Februar 2007. Frankfurt/Main

[13] **DSpace:** www.DSpace.org

[14a] **Liste der OPUS-Installationen des HBZ:** www.hbz-nrw.de/angebote/medienserver/opus/index_html#hostings [Stand 02.07.2007]

[14b] **Liste der OPUS-Installationen des Bayerischen OPUS-Servers:** www.opus-bayern.de/ [Stand 02.07.2007]

[14c] **Liste der OPUS-Hostings des KOBV:** www.kobv.de/deutsch/content/partner_info/opus.htm [Stand 02.07.2007]

[15] **Fedora:** www.fedora.info/

[16] **Forkel, Robert:** The Living Reviews Publishing Platform, Vortrag auf der Berlin V Tagung, Satellite Workshop: Services for Open Access Publishing, Golm, 31.03.2006. <http://dev.livingreviews.org/presentations/berlin4.html> [Stand 02.07.2007]

¹ Roar.EPrints.org, zuletzt gesehen am 2. Juli 2007

DIE VERFASSERIN

Susanne Dobratz ist Leiterin der AG Elektronisches Publizieren des Computer- und Medienservice und der Universitätsbibliothek der Humboldt-Universität zu Berlin, Universitätsbibliothek, Unter den Linden 6, 10099 Berlin, dobratz@cms.hu-berlin.de