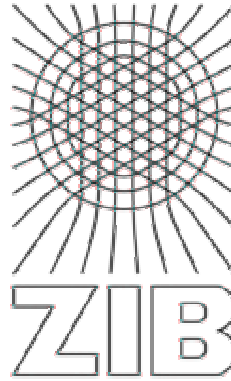

Konrad-Zuse-Zentrum
für Informationstechnik Berlin



Takustraße 7
D-14195 Berlin-Dahlem
Germany

Monika Kuberek

Open-Access –
Aufbau, Pflege und Betrieb lokaler Publikationsserver
in der KOBV-Zentrale

Gefördert
von der Senatsverwaltung für Wissenschaft, Forschung und Kultur des Landes Berlin,
vom Ministerium für Wissenschaft, Forschung und Kultur des Landes Brandenburg
und von den Mitgliedsbibliotheken des KOBV

Open Access – Aufbau, Pflege und Betrieb lokaler Publikationsserver in der KOBV-Zentrale

Monika Kuberek
(*kuberek@zib.de*)

Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)

ZIB Report 07-47

Dezember 2007

Abstract

Zur Unterstützung der Bibliotheken bei ihren Open-Access-Aktivitäten betreibt die KOBV-Zentrale seit Anfang 2005 den Service "Opus- und Archivierungsdienste". Die KOBV-Zentrale agiert als Application Service Provider (ASP) für sämtliche technischen Komponenten des Publikationsprozesses, indem sie die gesamte technische Infrastruktur bereitstellt und betreibt – angefangen bei den lokalen Publikationsservern bis hin zu lokalen Repositories zur Archivierung der elektronischen Dokumente.

Der vorliegende Text ist die schriftliche Fassung eines gleichnamigen Vortrages auf der 31. ASpB-Tagung "Kooperation versus Eigenprofil?" der Arbeitsgemeinschaft der Spezialbibliotheken, die vom 25.-28. September 2007 in der Technischen Universität Berlin stattfand.

Keywords: DINI-Zertifikat, Kooperativer Bibliotheksverbund Berlin-Brandenburg (KOBV), Open Access, Open Archive Initiative, Opus-Publikationssoftware

CR: H.0K, 4.0

Inhalt

| | |
|---|----------|
| Vorbemerkung | 3 |
| Application Service Providing (ASP) "Opus- & Archivierungsdienste" | 3 |
| Konsortialmodell des KOBV | 4 |
| Funktionen der Publikationssoftware Opus | 5 |
| Open Access / DINI-Zertifikat 2007 | 6 |

Vorbemerkung

Nicht zuletzt durch die Ansiedlung am Konrad-Zuse-Zentrum für Informationstechnik (ZIB) ist Open Access ein seit langem diskutiertes Thema in der Verbundzentrale des Kooperativen Bibliotheksverbundes Berlin-Brandenburg (kurz: KOBV-Zentrale).¹ Zur Unterstützung der Bibliotheken bei ihren Open-Access-Aktivitäten betreibt die KOBV-Zentrale seit Anfang 2005 den Service "Opus- und Archivierungsdienste".² Die Dienstleistung umfasst über den Betrieb des konsortial betriebenen Publikationswerkzeuges Opus hinaus auch die Archivierung der elektronischen Dokumente. Integraler Bestandteil des Dienstes sind weiterhin die Bereitstellung der technischen Infrastruktur³ und die Betreuung der Bibliotheken bei der Antragstellung für die DINI-Zertifizierung, Unterstützung der Bibliotheken als DataProvider im Sinne der Open Access Initiative und nicht zuletzt die Möglichkeit, den elektronischen Dokumenten eine dauerhafte Internet-Adresse zuzuweisen. Zusätzlich zu den in der Opus-Standardanwendung integrierten Recherchemöglichkeiten – Metadaten-Suche und Browsing – hat die KOBV-Zentrale sämtliche Metadaten und Volltexte mit Lucene indiziert und so mittels Open-Source-Suchmaschinentechnologie eine Volltextrecherche realisiert.

Application Service Providing (ASP) "Opus- & Archivierungsdienste"

Die KOBV-Zentrale agiert als Application Service Provider (ASP) für sämtliche technischen Komponenten des Publikationsprozesses, indem sie die gesamte technische Infrastruktur bereitstellt und betreibt – angefangen bei den lokalen Publikationsservern bis hin zu lokalen Repositories zur Archivierung der elektronischen Dokumente.

Die Serviceleistung umfasst auf Seiten der KOBV-Zentrale die Bereitstellung, Betrieb und Pflege eines performanten Servers mit hoher Ausfallsicherheit, so dass für die Bibliothek Beschaffung und Betrieb eines eigenen Servers entfallen. Auf dem Server ist die Publikationssoftware Opus sowie die notwendige Peripheriesoftware (z.B. PHP) installiert, deren Betrieb und Pflege ebenfalls die KOBV-Zentrale übernimmt. Zu diesem Zweck hat sie eine konsortiale Umgebung aufgebaut, in der sie für jede Bibliothek ihren lokalen Publikationsserver konfiguriert, d.h. eine eigene Opus-Instanz mit bibliotheksspezifischen Anpassungen einrichtet. Die Bibliothek braucht sich um Releases und Versions-Upgrades der Software nicht zu kümmern. Für einen neuen Teilnehmer an dem Service, der bereits einen eigenen Publikationsserver betrieben hatte, übernimmt die KOBV-Zentrale die Migration der Daten in die neue Umgebung.

Zur dauerhaften Adressierung der elektronischen Publikationen im Internet arbeitet die KOBV-Zentrale mit der Deutschen Nationalbibliothek zusammen, die einen URN-Resolving-Dienst⁴ für deutschsprachige Netzpublikationen aufgebaut hat. Die KOBV-Zentrale hat die seit Version 3.0 in Opus implementierte OAI-PMH-Schnittstelle in Betrieb genommen, ebenso die technischen Voraussetzungen für die DINI-Zertifizierung der lokalen Publikationsserver geschaffen – beides wesentliche Komponenten im Zusammenhang mit Open-Access; dazu weiter unten mehr.

Für die Speicherung und Archivierung der Publikationen stellt die KOBV-Zentrale den Bibliotheken Speicherplatz in bibliothekseigenen Repositories zur Verfügung. Dazu betreibt sie einen leistungsfähigen Server und sorgt für Sicherungs- und Back-up-Dienste sowie für eine hohe Ausfallsicherheit des Servers. Sowohl im Hinblick auf die Beschaffung von Speicherplatz als auch beim Einsatz modernster Sicherungstechnologie arbeitet die KOBV-Zentrale eng mit der IT-Abteilung des ZIB zusammen, kann auf die leistungsfähige Infrastruktur des ZIB zurückgreifen und damit Synergien nutzen, die aus der Anbindung der KOBV-Zentrale an das ZIB entstehen.⁵ Gemeinsam mit dem ZIB kümmert sich die

¹ So gehört der Vizepräsident des ZIB, Prof. Dr. Drs. h.c. Martin Grötschel, in dessen Abteilung die KOBV-Zentrale angesiedelt ist, zu den Initiatoren der "Berlin Declaration on Open Access ...", die 2003 im Rahmen einer Tagung der Max-Planck-Gesellschaft verabschiedet wurde.

² Der Dienst steht über den KOBV hinaus allen interessierten Bibliotheken offen. Leistungen und Konditionen sind auf den KOBV-Seiten ausführlich beschrieben, siehe: http://www.kobv.de/bib_opus_archivierung.html

³ Der technische Service der KOBV-Zentrale hat 2007 das Verfahren für das "DINI-Zertifikat 2007" erfolgreich durchlaufen.

⁴ Die URN (Uniform Resource Name) bezeichnet eindeutig den "Standort" eines Objektes im Internet und dient zur dauerhaften Adressierung und zuverlässigen Zitierfähigkeit von Online-Ressourcen. Mehr zum URN-Dienst der Deutschen Nationalbibliothek unter: http://www.d-nb.de/netzpub/erschl_lza/np_urn.htm

⁵ Das ZIB betreibt beispielsweise das Berliner Wissenschaftsnetz (<http://www.brain.de>) und ist – neben dem Regionalen Rechenzentrum Niedersachsen der Universität Hannover – einer der beiden Standorte des Super-

KOBV-Zentrale auch um die Langzeitarchivierung der Daten und verfolgt intensiv die Projekte der DNB hinsichtlich der Langzeitverfügbarkeit.

Innerhalb dieses Konzeptes braucht sich die Bibliothek selbst lediglich um die inhaltlich-fachlichen Arbeiten des Publikationsprozesses zu kümmern, wie Betreuung der Autoren, Ausarbeitung der bibliotheksspezifischen Leitlinien für das Publizieren, Anpassung der Hilfetexte in Opus sowie der Webseiten des Publikationsservers. Auch die Qualitätskontrolle der von den Autoren eingegebenen Metadaten sowie die Freigabe der elektronischen Dokumente im Internet liegt in der Hand der Bibliothek. Ebenso entscheidet die Bibliothek über den Nachweis der Publikationen, wobei sie im Rahmen des KOBV-Services Ermessensspielräume hat, die sie ganz oder teilweise ausschöpfen kann. So liegt es in der Entscheidung der Bibliothek, ob sie sich am URN-Dienst der Deutschen Nationalbibliothek beteiligt⁶ und ob sie ihren Publikationsserver für das Abrufen der Metadaten durch OAI-ServiceProvider freigibt. In allen Fällen wird sie von der KOBV-Zentrale durch Datenlieferungen oder Bereitstellung von Schnittstellen unterstützt. Für die Bibliothek liegt der große Vorteil des KOBV-Dienstes darin, dass sie die gesamte Technik an die KOBV-Zentrale outsourcen kann und selbst keinerlei technisches Kenntnisse mitbringen oder vorhalten muss, um ihren Publikationsserver zu betreiben.

Konsortialmodell des KOBV

Die KOBV-Zentrale betreibt die Serviceleistung in einer konsortialen Umgebung. In dieser Umgebung hat jede Bibliothek eine eigene Opus-Instanz mit einem spezifischen, Passwort-geschützten Zugriff auf ihren Admin-Bereich sowie ein eigenes Repository zum Archivieren der bibliothekseigenen Dokumente.

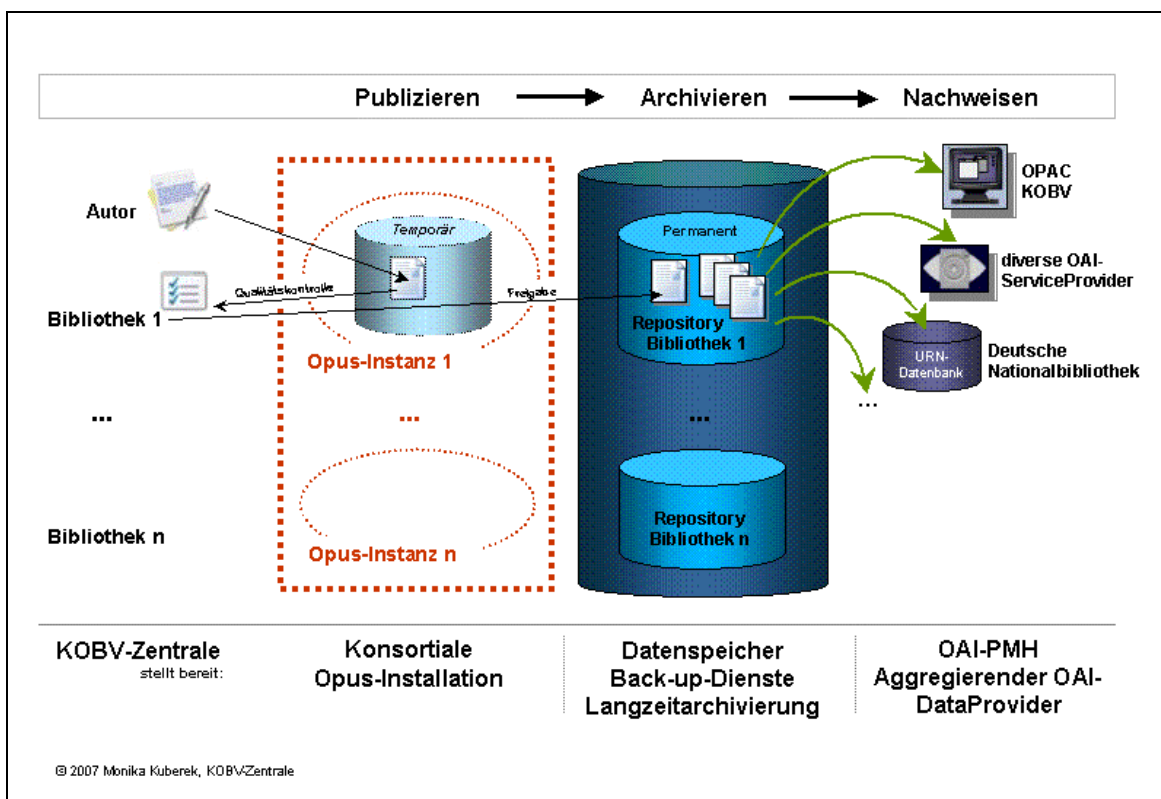


Abb. "Konsortiales Publizieren und Archivieren": Im Konsortialmodell des KOBV werden Aufbau und Betrieb der gesamten technischen Infrastruktur – Server, Software, Datenspeicher, Schnittstellen – von der KOBV-Zentrale übernommen. Die Bibliothek braucht sich lediglich um die bibliothekarisch-fachlichen Arbeiten des Publikationsprozesses zu kümmern.

computers, mit dem sechs Bundesländer gemeinsam den Norddeutschen Bund zur Förderung des Hoch- und Höchstleistungsrechnens betreiben (<http://www.hlrn.de/>).

⁶ obligatorisch für Dissertationen

Nach diesem Modell betreibt die KOBV-Zentrale derzeit 16 Publikationsserver, davon 14 im Produktivbetrieb.⁷

Eingesetzt wird Opus in der Standardversion, derzeit Version 3.0. Die Migration auf eine nächsthöhere Version erfolgt für sämtliche lokalen Instanzen durch die KOBV-Zentrale in Absprache mit den Bibliotheken zeitnah zur Freigabe der Version. Selbstverständlich hat auch in der konsortialen Umgebung jede Bibliothek die Möglichkeit, ihre lokale Instanz in ihren Web-Auftritt und das Bibliotheks- bzw. Universitäts-eigene Design einzupassen. Diese Integration erfolgt in den Bibliotheken mehr oder weniger intensiv. Während einigen Bibliotheken das Einbinden von Bibliotheks-Logo und -Namen in die Opus-Standardoberfläche genügt, gehen andere sehr viel weiter, so dass gegebenenfalls nur noch "für Eingeweihte" die Opus-Herkunft des lokalen Publikationsservers erkennbar ist.⁸

Auch das Archivieren funktioniert nach einem konsortialen Modell, bei der jede Bibliothek einen eigenen Speicherbereich auf dem KOBV-Server erhält. Das bibliothekseigene Repository umfasst als Grundausstattung einen Speicherplatz von 100 GigaByte und kann – dem Bedarf der Bibliothek entsprechend – jederzeit erweitert werden.

Funktionen der Publikationssoftware Opus

Der einfach zu handhabende Workflow – sowohl für die Autoren als auch für die Bibliotheken – waren wesentliche Gründe dafür, dass sich Opus in der Region Berlin-Brandenburg als Publikationssoftware durchgesetzt hat. Der Opus-Workflow sieht vor, dass die Autoren ihre Metadaten selbst eingeben und ihre elektronischen Dokumente selbst hochladen, zunächst in ein temporäres Verzeichnis. In der anschließenden Qualitätskontrolle überprüft die Bibliothek, ob Metadaten und Dokumente funktional in Ordnung sind und den bibliothekseigenen Konventionen der Veröffentlichung entsprechen. Ist dies der Fall, werden die Dokumente – es genügt ein Mausklick – von der Bibliothek in ihr Repository "geschoben", wo sie permanent gespeichert werden. Je nach den spezifischen Zugriffsmodalitäten (s.u.) sind nun die Metadaten samt den zugehörigen Volltexten oder auch nur die Metadaten für den Zugriff im Internet bzw. über das OAI-Protokoll freigegeben.

Den Bibliotheken bietet Opus ab Version 3 neben der Verwaltung von Einzelpublikationen auch die Zusammenfassung und Verwaltung der elektronischen Dokumente in Kollektionen und Schriftenreihen. Die Veröffentlichungsart "Kollektion" kann die Bibliothek dazu einsetzen, Sammlungen abzubilden; diese können beliebig hierarchisch "verschachtelt" werden. Bei Schriftenreihen bildet Opus die Verknüpfungen von Dokumenten einer Reihe ab und versieht die Dokumente automatisch mit einer fortlaufenden Bandnummer.

Ab Version 3 bietet Opus den Bibliotheken weitere Funktionalitäten, die die Verwaltung der Dokumente erleichtern. So können neben den in Opus üblichen pdf-Dateien nunmehr auch die zugrundeliegenden Originaldateien in einem beliebigen Format hochgeladen werden – beispielsweise von Dissertationen, deren Aufbewahrung in vielen Prüfungsordnungen vorgeschrieben ist. Original- und Publikationsdateien können damit in einem Arbeitsgang im Bibliotheksrepository abgelegt und gespeichert werden. Weiterhin lassen sich Metadaten nun auch hochladen, ohne dass das Dokument vorhanden ist – eine Forderung aus der wissenschaftlichen Community insbesondere im Hinblick auf Preprint-Server, wo häufig die Zitierfähigkeit schon zu einem sehr frühen Zeitpunkt gegeben sein soll, auch wenn das Dokument selbst möglicherweise noch nicht das Veröffentlichungsstadium erreicht hat. Sobald das Dokument vorliegt, kann es über den implementierten Dateimanager von der Bibliothek auf komfortable Weise hochgeladen und im Bibliotheksrepository abgelegt werden.

Die in Opus implementierte Standardrecherche – das Browsen nach Dokumententypen, Schriftenreihen, Kollektionen sowie die Suche nach bibliographischen Metadaten – wurde von der KOBV-Zentrale durch eine Volltextrecherche erweitert: Dazu werden Metadaten und sämtliche Volltexte mit Lucene⁹ indiziert und suchbar gemacht. Auf diese Weise hat die KOBV-Zentrale auch für die Opus-Anwender

⁷ Die konsortial betriebenen lokalen Publikationsserver sind auf den KOBV-Seiten aufgelistet unter: http://www.kobv.de/bib_opus_teilnehmer .

⁸ Als Beispiel für die sehr weitgehende Integration sei hier auf den Publikationsserver der UB der Universität Potsdam hingewiesen: <http://digibib.kobv.de/ubp> .

⁹ Lucene ist eine Open-Source-Programmbibliothek und Lucene-Anwendungen sind auch in anderen Produkten des KOBV im Einsatz, z.B. im KOBV-Volltextserver (<http://volltexte.kobv.de>) oder in der von der KOBV-Zentrale entwickelten und betriebenen Lucene-Anwendung für Bibliotheken, die im Wissenschaftspark Albert Einstein Potsdam im Routinebetrieb läuft (<http://waesearch.kobv.de>) .

die erweiterten Suchmöglichkeiten und Vorteile der Suchmaschinenteknologie – Rechercheergebnisse in Sekundenbruchteilen, Unschärfe-(Fuzzy)-Suche, Meinten-Sie-Funktionalität, Ranking, ... – nutzbar gemacht.

Nicht alle Bibliotheken möchten – oder können – die Freigabe sämtlicher Volltexte ihres Publikations-servers im Internet gestatten; z.B. können Urheberrechtsgründe dagegen sprechen. Hier bietet Opus differenzierte Zugriffsbeschränkungen, auch für einzelne Publikationen, und erlaubt der Bibliothek auf diese Weise ein Dokumenten-spezifisches Veröffentlichungsmanagement: Der freie Zugang auf alle Dokumente ist ebenso möglich wie die Beschränkung des Zugangs auf den Campusbereich der Hochschule oder die vollständige Sperrung des Volltextes für jeden Zugriff von außen. In all diesen Fällen bleiben Metadaten und Abstracts suchbar und sind über das OAI-Protokoll abrufbar – der Nachweis der elektronischen Publikationen ist mithin immer gewährleistet und entspricht damit den von der Open Archive Initiative aufgestellten Konventionen des freien Zugriffs auf elektronische Publikationen.¹⁰

Open Access / DINI-Zertifikat 2007

Mit der Implementierung der OAI-PMH-Schnittstelle erfüllt Opus eine wesentliche Anforderung der Open Archive Initiative: die weltweite freie Verfügbarkeit der Metadaten über ein offenes Standardprotokoll. In Opus 3 ist die OAI-PMH-Schnittstelle Version 2 implementiert. Die KOBV-Zentrale hat die Schnittstelle aktiviert und damit für jede Bibliothek, die am Konsortialmodell teilnimmt, die Möglichkeit geschaffen, als DataProvider zu agieren und die Metadaten ihrer Publikationen weltweit von den diversen OAI-ServiceProvidern abrufen zu lassen. Auf diese Weise kann die Bibliothek ihre elektronischen Publikationen einer internationalen Fach-Community zur Verfügung stellen. Sowohl die Schnittstelle als auch der freie Zugriff auf alle Metadaten sind für die Bibliotheken wichtige technische Schritte im Zuge einer Open-Access-Lösung – neben den ebenso grundlegenden organisatorischen Voraussetzungen innerhalb der Hochschulen, die oftmals sehr viel langwieriger zu realisieren sind.

Die OAI-PMH-Schnittstelle ist Voraussetzung für die DINI-Zertifizierung. Opus 3.0 – in Kombination mit Tools zur Erstellung von Zugriffsstatistiken, die zwar nicht in Opus implementiert sind, aber von der KOBV-Zentrale entwickelt und zur Verfügung gestellt werden – erfüllt die technischen Voraussetzungen für das „DINI-Zertifikat Dokumenten- und Publikationsservice 2007“¹¹, einem in der Open-Access-Bewegung hoch angesehenen Qualitätssiegel. Das DINI-Zertifikat 2007, das im Februar 2007 freigegeben wurde, trägt aktuellen internationalen Entwicklungen Rechnung und beschreibt, wie Dokumenten- und Publikationsservices das Open-Access-Publizieren wissenschaftlicher Dokumente unterstützen und wie sie sich als Institutional Repositories in ihrer jeweiligen Institution positionieren können.

Die Teilnehmer am Konsortialmodell des KOBV können ihre Anwendung durch DINI zertifizieren lassen. Die Zertifizierung ist optional und kann von jeder Einzelbibliothek beantragt und erworben werden. Im September 2007 hat die UB der Universität Potsdam als erste KOBV-Bibliothek (und zweite Bibliothek bundesweit) das Dini-Zertifikat 2007 erhalten; dabei hat der technische Service der KOBV-Zentrale das Zertifizierungsverfahren erfolgreich durchlaufen. Von den Gutachtern ausdrücklich positiv hervorgehoben wurde die arbeitsteilige Konzeption im KOBV, in der die KOBV-Zentrale für die technischen Belange zuständig ist und die Bibliotheken sich nur noch um bibliothekarisch-fachliche Dinge kümmern müssen.

In der konsortialen Umgebung des KOBV sind sowohl die Bibliothek als auch die KOBV-Zentrale an der DINI-Zertifizierung beteiligt: Begutachtet werden der jeweilige organisatorische Workflow – d.h. die bibliothekseigene Komponente – sowie die technische Infrastruktur – d.h. die Servicekomponente der KOBV-Zentrale. Diese Servicekomponente ist im Konsortialmodell für alle lokalen Publikationsserver identisch. Die KOBV-Zentrale unterstützt die Bibliotheken bei der Antragstellung, stellt ihnen für die Beantwortung der technischen Fragen den ausgefüllten technischen Fragebogenteil (der bereits positiv begutachtet wurde) zur Verfügung. und betreut sie während des gesamten Prozesses der Zertifizierung.

¹⁰ Zur Open Archive Initiative und zu OAI-PMH (Open Archives Initiatives Protocol for Metadata Harvesting) siehe: <http://www.openarchives.org/>

¹¹ DINI = Deutsche Initiative für Netzwerkinformation e.V. (<http://www.dini.de>)