

Qualitätskontrolle für Dokumenten- und Publikationsserver:

Das DINI Zertifikat¹

Susanne Dobratz, Berlin

Das DINI- Zertifikat für Dokumenten- und Publikationsserver führt durch die Definition von Mindeststandards und Empfehlungen erstmals eine Qualitätskontrolle ein.

For the first time the DINI Certificate for Document and Publication Repositories introduces a quality control for servers by defining minimum standards and recommendations.

Dokumentenserver in Deutschland

Seit 1996/97 werden in Deutschland an den Hochschulen Dokumentenserver aufgebaut, die elektronische Publikationen aufnehmen. Autoren werden dabei oft geworben, indem sie auf die Vorteile hingewiesen werden, die diese neue Möglichkeit der Publizierens bringt: eine digitale Publikation ist schneller verfügbar, preiswert, für Bibliotheken und Nutzer Platz sparend und weltweit verfügbar. Das bringt dann Zweifler dazu, ebenso auf die Nachteile hinzuweisen: digitale Publikationen gelten als flüchtig, anfällig für Veränderungen und Missbrauch, beliebig veränderbar und nicht zitierbar.

Bibliotheken und Institutionen, die sich also der Aufgabe stellen, einen Dokumentenserver zu etablieren, müssen Lösungen finden, um die erwähnten Nachteile auszuräumen. Hierzu ist es notwendig, bestimmte Standards, Verfahren und Technologien einzusetzen, die in Analogie zum gedruckten Buch die Qualität und Stabilität des Inhalts gewährleisten, die Dauerhaftigkeit und Verfügbarkeit der Publikation garantieren, ebenso deren Sichtbarkeit und Auffindbarkeit und die Autoren und Nutzern eine gewisse Sicherheit im Umgang mit den digitalen Dokumenten garantieren. Um den klassischen Sammelauftrag von Bibliotheken zu erweitern, müssen sich die Bibliotheken zu digitalen Bibliotheken entwickeln.

Inzwischen gibt es an fast jeder deutschen Hochschule einen Dokumentenserver. Die Deutsche Bibliothek sammelt seit 1998 die digitalen Dissertationen der Hochschulen und stellt mit dem DissOnline-Portal² eine entsprechende Unterstützung für Autoren, Institute, Bibliotheken und Universitäten zur Verfügung.

Dennoch haben sich in den letzten Jahren eine Vielzahl von Lösungen entwickelt, die mitunter nicht miteinander kompatibel sind. In einer Umfrage³, die die DINI- Arbeitsgruppe „Elektronisches Publizieren“ gemeinsam mit dem DFG-Projekt Proprint⁴ 2002/2003 durchgeführt hat, wurde deutlich, dass die dringende Notwendigkeit besteht, Dokumentenserver zu vereinheitlichen, um die Einhaltung nationaler und internationaler Standards, wie Dublin Core Metadata Set oder OAI, zu garantieren, die Benutzung proprietärer Dateiformate zu vermeiden, die Authentizität und Integrität der Dokumente und des Dokumentenservers zu bewahren, sowie die langfristige Verfügbarkeit digitaler Dokumente zu sichern.

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

Das Entstehen der Dokumentenserver ist eine der Entwicklungen, die die Deutsche Initiative für Netzwerkinformation begleitet. Vorgängeraktivitäten, wie die Formulierung eines Grundsatzpapiers mit dem Titel „Information und Kommunikation in lokalen, nationalen und internationalen Datennetzen – Memorandum zur Notwendigkeit einer nationalen Koordination“ gab es seit 1991. DINI wurde von der Deutschen Forschungsgemeinschaft (DFG) von 1999 bis 2001 als Projekt gefördert. Ziel ist es, Absprachen, Kooperationen, Empfehlungen und Standards, die durch den Wandel in der Kommunikationstechnologie innerhalb der Informations-Infrastrukturen der Hochschulen und anderer Forschungseinrichtungen notwendig sind, zu fördern und zu initiieren. Dabei legt DINI Wert darauf, daß die Verbesserung der Informations- und Kommunikationsdienstleistungen und die dafür notwendige Entwicklung der



Informations-Infrastrukturen an den Hochschulen und Fachgesellschaften regional und überregional gefördert werden.

Durch Absprachen und Arbeitsteilung zwischen den Infrastruktureinrichtungen soll das Dienstleistungsangebot weiter verbessert werden. Über die Zusammenarbeit der Infrastruktureinrichtungen hinaus ist es Ziel von DINI die gemeinsame Entwicklung von Standards und Empfehlungen zu fördern.

Seit 18. März 2002 ist DINI ein eingetragener Verein mit einem Hauptausschuss, einem Vorstand mit 2 Vorstandsprechern und Vertreter aus 4 Bereichen: Bibliotheken, Medienzentren, Rechenzentren und der Wissenschaft.

In DINI diskutieren Arbeitsgruppen unterschiedliche Fragestellungen der IUK-Technologie. Eine dieser AG's ist die Arbeitsgruppe Elektronisches Publizieren (E-Pub), die bereits 2002 Empfehlungen⁵ zum Elektronischen Publizieren an Hochschulen formuliert hat, mit dem Ziel Standards und Technologien in diesem Bereich in Deutschland besser zu verbreiten. 2003 wurde auf der Basis dieser Empfehlungen das DINI- Zertifikat für Dokumenten- und Publikationsserver erarbeitet.

DINI- Zertifikat: Mindeststandards und Empfehlungen

Ziel des DINI- Zertifikats für Dokumenten- und Publikationsserver ist es, eine detaillierte Beschreibung der Anforderungen an einen Dokumenten- und Publikationsserver vorzulegen, Entwicklungsrichtungen bei der Gestaltung von Servern und beim Austausch von Informationen über diese Server aufzuzeigen sowie eine für Nutzende und Betreiber sichtbare Dokumentation der Einhaltung von Standards und Empfeh-

lungen durch diesen Server zu generieren.

Mit der Vergabe eines Zertifikats ermöglicht DINI e.V. erstmals eine Qualitätskontrolle für Dokumenten- und Publikationsserver. Um ein Zertifikat zu erhalten, sind Mindestanforderungen an einen Server und seine Betreiber gestellt, die zur Wahrung einer modernen wissenschaftlichen Kommunikation unabdingbar sind. Gleichzeitig werden Empfehlungen gegeben, die die heute vorhersehbare Entwicklung charakterisieren und damit künftige Serveranforderungen darstellen.

Leitlinien (Policy)

Als Mindeststandard wird verlangt, dass der Betreiber des Dokumentenservers Leitlinien für inhaltliche Kriterien sowie den Betrieb des Servers formuliert und diese veröffentlicht. Darin müssen die Rechte und Pflichten des Betreibers des Dokumentenservers sowie die der Autoren/Herausgeber festgeschrieben sein. Weiterhin muss eine Policy Festlegungen zur inhaltlichen, funktionalen und technischen Qualität der Dokumente enthalten, die auf dem Dokumentenserver veröffentlicht werden. Es müssen Garantien für bestimmte Archivierungszeiträume in Abhängigkeit der inhaltlichen, funktionalen und technischen Qualität der Dokumente festgelegt werden sowie Verfahren, die der Betreiber zum Betrieb des Dokumentenservers nutzt und bestimmter Leistungen, die der Betreiber des Dokumentenservers den Autoren und Herausgebern anbietet.

Autorenbetreuung

Als Mindeststandard für Dokumentenserver gilt hier das Vorhandensein eines nach außen sichtbares Beratungsangebotes (z.B. WWW-Seiten und E-Mail-Kontaktadressen, telefonische und persönliche Autorenbetreuung) und die Unterstützung des gesamten Publikationsprozesses (einschließlich rechtlicher und technischer Problemstellungen) durch den Betreiber.

Empfohlen wird, Kurse zum elektronischen Publizieren mindestens einmal im Semester anzubieten sowie Angebote zum strukturierten Schreiben für Autoren bereitzustellen. Die Verwendung von Helpdesk-Systemen wird als eine zukünftig wichtige Entwicklung angesehen und empfohlen.

Rechtliche Aspekte

Bestimmte rechtliche Gegebenheiten müssen für den Dokumentenserver

mindestens geklärt sein, diese sind: die Berechtigung für den Betreiber, die eingereichten Dokumente auf dem Server zu veröffentlichen (Recht zur elektronischen Speicherung, insbesondere in Datenbanken, und zum Verfügbarmachen für die Öffentlichkeit zum individuellen Abruf, zur Wiedergabe auf dem Bildschirm und zum Ausdruck beim Nutzer [Online-Nutzung], auch auszugsweise), die Berechtigung die Dokumente an eine Archivierungsinstitution weiter zu geben und das Recht, die Dokumente durch den Betreiber zum Zweck der Sicherung der Langzeitverfügbarkeit technisch verändern zu können. Desweiteren muss ein Haftungsausschluss in einem Disclaimer formuliert sein.

Als Empfehlung wird die Verankerung der unterschiedlichen Rechte in den Metadaten ausgesprochen.

Authentizität und Integrität

Bei der Garantie der Authentizität und Integrität gibt es zwei Aspekte. Die ist zum einen die Sicherheit des Servers. Hier wird als Mindeststandard das Vorliegen einer Dokumentation des technischen Systems angesehen und es wird davon ausgegangen, dass der Server darauf ausgelegt ist, 365 Tage im Jahr zur Verfügung zu stehen. Es muss ein System zur Sicherung der Server, Metadaten und Dokumente verfügbar sein und die Aufnahme von Dokumenten aus technischer Sicht muss kontrollierbar und nachweisbar sein.

Empfohlen wird die Nutzung einer SSL – Zertifizierung für den Server, um sichere Übertragungen z.B. über WWW-Schnittstellen bereitstellen zu können.

Zum anderen werden Anforderungen zur Gewährleistung der Sicherheit der Dokumente gestellt. Hier gilt als Mindeststandard die Verwendung von Persistent Identifier, wie zum Beispiel URNs, wie sie durch Die Deutsche Bibliothek vergeben werden⁶. Dabei gilt als Regel: Ein inhaltlich verändertes Dokument ist wie ein neues Dokument zu behandeln. Pflicht ist die Archivierung der eingereichten Dateien des Autors im Ablieferungsformat.

Folgende Empfehlungen werden gegeben: Es sollte ein Verfahren zum Nachweis der Unversehrtheit der Dokumente (z.B. Hash-Wert) und fortgeschrittene digitale Signatur nach § 2 Abs. 2 SigG 2001 zum Einsatz kommen. Weiterhin wird empfohlen, Archivierungsformate aus den gelieferten Dokumentformaten zu erzeugen, um diese dann in

Langzeitarchivierungssysteme exportieren zu können. Die Darstellung der Dokumente sollte in einem gängigen Präsentationsformat geschehen. Wenn dies nicht möglich ist, sollte ein Viewer entweder zum Herunterladen oder eine Internetreferenz für den Viewer angeboten werden.

Erschließung

Hier gibt es Anforderungen an die Sacherschließung, an den Export von Metadaten und an bereitzustellende Schnittstellen.

Mindeststandards sind das Vorhandensein einer definierter Policy zur Sacherschließung, die Durchführung einer verbalen Sacherschließung durch freie Schlagwörter oder klassifikatorische Erschließung, die Gewährleistung der freien Zugänglichkeit der Metadaten und die Nutzung von Dublin Core Simple (ANSI/NISO Z39.85-2001)⁷, sowie die Bereitstellung einer OAI2.0- Schnittstelle und einer Webserverschnittstelle für Nutzer.

Empfohlen wird die Anwendung der Dewey-Dezimalklassifikation (DDC) gemäß der Verwendung in der Deutschen Nationalbibliografie als allgemeine klassifikatorische Erschließung aller Dokumente und eines weiteren normierten Systems verbaler oder klassifikatorischer Erschließung (z.B. Schlagwortnormdatei, LoC, CCS, MSC, PACS etc.) sowie die Vergabe englischsprachiger Schlagwörter und die Bereitstellung eines deutschen und eines englischsprachigen Abstracts.

Für den Metadatenexport empfiehlt DINI die Nutzung von Dublin Core qualified oder von ONIX- Daten⁸, das Anbieten weiterer technische und/oder Archivierungsmetadaten (z.B. METS⁹), bzw. Print on demand Daten. Ebenso wird der Export in Langzeitarchivierungssysteme und bibliographische Datenbanken (z. B. Literaturverwaltungsprogramme, Verbünde, OPACs) empfohlen. Dazu sollten Schnittstellen bereit gestellt werden, die z.B. SOAP¹⁰ als Webserviceschnittstelle, Z39.50¹¹ oder auch Erweiterungen des OAI-Protokolls¹² für den Austausch komplexerer Metadatenschemata nutzen.

Zugriffsstatistik

Jeder Dokumentenserver muss im Rahmen der rechtlichen Bestimmungen mindestens eine eigene konsistente Zugriffsstatistik führen (Webserver-Logs).

Empfohlen wird, die Statistik als dynamisches Metadatum dem jeweiligen Dokument zuzuordnen.

Langzeitverfügbarkeit

Um die Dokumente des Dokumentenservers später für eine Langzeitarchivierung nutzen zu können, muss es eine dauerhafte Verbindung der Metadaten mit den Dokumenten (z.B. Verbindung Über Persistent Identifier oder zusammen in einem Container) geben. Weiterhin ist eine Mindestzeit der Dokumentverfügbarkeit in der Policy, die 5 Jahre nicht unterschreiten darf, zu definieren und umzusetzen.

Die Sicherstellung der Langzeitverfügbarkeit durch ggf. Kooperation mit einer Archivierungsinstitution wird dringend empfohlen, z.B. im Rahmen von nestor¹³, der Kooperationsplattform Langzeitarchivierung.

Vergabemodus

Vergeben wird das DINI- Zertifikat für Dokumenten- und Publikationsserver durch die DINI-Geschäftsstelle bzw. durch eine von ihr beauftragte Arbeitsgruppe gegen eine Gebühr. Voraussetzung ist die Antragstellung über die DINI- Webseite¹⁴. Verbunden mit dem Zertifikat, das die entsprechende Jahres-

zahl des Ausstellungsjahres enthält, wird eine Urkunde die Bescheinigung über die Erfüllung der Mindestanforderungen für Dokumenten- und Publikationsserver ausgestellt.

Ausblick

Das Zertifikat wird regelmäßig an aktuelle technische und organisatorische Rahmenbedingungen angepasst. Es ist nach Ausstellung zeitlich nicht befristet. Aufgrund des eingetragenen Ausstellungsjahres ist jedoch jederzeit überprüfbar, welchen Leistungsanforderungen der zertifizierte Server bzw. Teile des Servers genügen.

Weitere Erläuterungen zu den einzelnen Punkten sowie Beispiele können aus der DINI- Schrift „DINI- Zertifikat für Dokumenten- und Publikationsserver“¹⁵ entnommen werden.

.....

¹ <http://www.dini.de/dini/zertifikat/zertifikat2.php>

² <http://www.dissonline.de/>

³ siehe http://www.dini.de/fragebogen/tvqval.php4?my_initglobal=digdoc/digdocinitglobal.inc

⁴ <http://edoc.hu-berlin.de/proprint>

⁵ <http://www.dini.de/documents/DINI-EPUB-Empfehlungen-2002-03-10.pdf>

⁶ siehe <http://www.persistent-identifier.de>

⁷ http://www.niso.org/standards/standard_detail.cfm?std_id=725

⁸ <http://www.editeur.org/onix.html>

⁹ <http://www.loc.gov/standards/mets/>

¹⁰ <http://www.w3.org/TR/SOAP/>

¹¹ http://www.niso.org/standards/resources/Z3950_Resources.html

¹² <http://www.openarchives.org/OAI/openarchivesprotocol.html>

¹³ <http://www.langzeitarchivierung.de>

¹⁴ <http://www.dini.de/dini/zertifikat/fragebogen.php>

¹⁵ <http://www.dini.de/documents/Zertifikat.pdf>

Diplom-Informatikerin Susanne Dobratz
Leiterin der Arbeitsgruppe
Elektronisches Publizieren

Humboldt- Universität zu Berlin
Universitätsbibliothek/Computer- und
Medienservice
Arbeitsgruppe Elektronisches Publizieren
Erwin Schrödinger- Zentrum
Rudower Chaussee 26
12489 Berlin
dobratz@cms.hu-berlin.de