

**Ute Schwens**

**Sammlung und Bereitstellung elektronischer Dissertationen – die Praxis Der Deutschen Bibliothek**

**Ablieferung elektronischer Publikationen**

Wie auf alle Bibliotheken, so kommt in der letzten Zeit auch auf Die Deutsche Bibliothek eine neue Publikationsform zu, die in der Handhabung und Archivierung völlig anderen Gesetzen unterworfen ist, als die bisherigen Formen – die elektronischen Publikationen.

Darunter verstehen wir

- Offline-Publikationen auf beliebigen Datenträgern (CD-ROM, Disketten u.a.)
- Online-Publikationen, die entweder
  - durch Digitalisierung gedruckter Dokumente entstehen oder
  - originär digital vorliegen, wie z.B. Online-Dissertationen oder Netzpublikationen von Verlagen oder sonstigen Produzenten.

Die Deutsche Bibliothek hat einen gesetzlichen Auftrag zur vollständigen Sammlung, Erschließung und dauerhaften Sicherung deutschsprachiger und in Deutschland erschienener Publikationen. Das Gesetz bezieht auch elektronische Publikationen ein, sofern sie auf physischen Datenträgern verbreitet werden.

Die Ablieferung von Netzpublikationen ist bisher im Gesetz über Die Deutsche Bibliothek nicht vorgesehen. Die Deutsche Bibliothek hält auch eine sofortige Gesetzesnovellierung nicht für sinnvoll, sondern sucht statt dessen das Gespräch mit Verlagen und Universitäten, um gesicherte Erkenntnisse hinsichtlich technischer Erfordernisse und einem abgestimmten Handeln aller Beteiligten zu sammeln, die in eine spätere gesetzliche Regelung einfließen können.

Auf diese Zusammenarbeit mit den Verlagen wird im Nachspann noch einmal Bezug genommen.

Im folgenden geht es zunächst einmal nur um die Online-Dissertationen.

## Online-Dissertationen

Der Geschäftsgang für Online-Dissertationen sieht im Überblick wie folgt aus:

Anmeldeformular	Elektronische Dissertationen und Habilitationen werden in Der Deutschen Bibliothek seit 1. Juli 1998 gesammelt. Der Kontakt zwischen der Universität und Der Deutschen Bibliothek erfolgt dabei
↓	über die Universitätsbibliothek. Diese meldet das Vorliegen der
Email	Netzdissertation per Email über ein auf der Website liegendes
↓	HTML-Formular an. Weitere Möglichkeiten der Meldung von
	Dissertationen werden ab Anfang 1999 angeboten, z.B. die direkte
	Lieferung von HTML-Meta-tags per Email oder als Email-Anhang.
Pull	Nach Empfang der Meldung einer Dissertation übernimmt Die
↓	Deutsche Bibliothek das Dokument selbst vom jeweiligen
deposit.ddb.de	Universitätsserver (pull) und legt es auf dem eigenen
	Dokumentserver deposit.ddb.de ab.
ILTIS	Die Dokumentbearbeitung folgt einem elektronischen Laufzettel:
↓	Formalerschließung auf der Grundlage von RAK-NBM,
(DNB, Reihe H)	Sacherschließung nach den RSWK, Anzeige der Netzdissertation in
↓	der Reihe H der Deutschen Nationalbibliographie. Die
	bibliographischen Angaben werden über die Zentralen
	bibliographischen Dienstleistungen vertrieben.
OPAC / Gateway	Der Zugriff auf die elektronischen Dokumente ist sowohl über den
	OPAC als auch über das Z39.50-Gateway möglich. Ein Zugriff direkt
	über den Dokumentserver ist geplant.

### ↳ Template und Metadaten

Im Hinblick auf das im Internet verwendete Template arbeitet Die Deutsche Bibliothek eng mit der Metadatengruppe des DFG-Projekts ‚Dissertationen online‘ zusammen.

Ziel dieser Zusammenarbeit ist es, den Universitäten und Universitätsbibliotheken eine Möglichkeit zu geben, **ein** Formular für den Gang der elektronischen Dissertation von der Abgabe an den Fachbereich durch den Promovenden bis zur Abgabe an Die Deutsche Bibliothek durch die Universitätsbibliothek zu nutzen.

Die einzelnen Felder des Templates sind

a) inhaltlich und

b) hinsichtlich ihrer Zuordnung zu Dublin Core Elementen

zwischen Der Deutschen Bibliothek und der Metadatengruppe ‚Dissertationen online‘  
abgesprochen. Ein hinter dem Formular liegendes CGI-Skript setzt die Angaben in eine  
Email mit HTML-Meta-tags um, die die abgesprochene Dublin Core-Datenstruktur abbildet.

Diese Formatstruktur für den Dublin Core-Datensatz wird demnächst auf den einzelnen  
Websites zu finden sein.

Auf der Basis dieser Absprache plant Die Deutsche Bibliothek, zusätzlich zu der Nutzung  
des Anmeldeformulars im Internet eine Dublin Core Schnittstelle anzubieten. Das bedeutet,  
daß Universitätsbibliotheken wählen können, ob sie unser Template oder den MyMetaMaker  
for Theses aus dem Projekt Dissertationen Online für die Meldung einer Dissertation nutzen  
oder ob sie an Die Deutsche Bibliothek einen Dublin Core Datensatz per Email schicken. Die  
Bedingungen für die Schnittstelle werden ca. Anfang 1999 über die Websites u.a.  
bekanntgegeben.

#### ↳ Geschäftsgang in Der Deutschen Bibliothek

Derzeit werden die Angaben aus dem Anmeldeformular beim Eingang in Die Deutsche  
Bibliothek über ein Umsetzungsskript in eine E-mail umgewandelt, die sozusagen den  
elektronischen Laufzettel der Dissertation beim ‚Gang durch das Haus‘ darstellt.

Dieser virtuelle Gang des Dokumentes durch das Haus zieht in den einzelnen Abteilungen  
folgende Aktivitäten nach sich:

Abt. 3,  
Erwerbung

Kurzkatalogisat in ILTIS,  
Virencheck, Authentizitätsschutz (Fingerprint),  
Erstellen einer Dokumentbeschreibungseite in HTML für die  
Kurzdarstellung auf dem  
Dokumentenserver, der sog. Frontpage,  
Ablegen des Dokumentes auf dem Server.

Abt. 4,  
Formalerschließung

Die formale Erschließung der Online-Publikationen erfolgt auf der  
Grundlage von RAK-NBM im PICA-ILTIS-Format, das für  
elektronische Off- und Online-Publikationen erweitert wurde und  
seit Ende Mai 1998 zum Einsatz kommt. Die  
Erschließungskonventionen haben einen vorläufigen Charakter, da  
zu einem späteren Zeitpunkt die Ergebnisse derzeit laufender  
Projekte gegebenenfalls zu anderen Verfahrensweisen führen  
werden.  
Besondere Berücksichtigung finden hierbei die Ergebnisse des  
Projekttes META-LIB.

Abt. 5,  
Sacherschließung

Sachkatalogisierung nach RSWK unter Einbeziehung der im  
Anmeldeformular mitgelieferten Erschließungshinweise.

Nach Ablauf des Geschäftsgangs erscheint die bibliographische Anzeige des elektronischen Dokuments in Reihe H der Deutschen Nationalbibliographie. Die Publikation ist über ILTIS, OPAC oder Gateway recherchierbar, Dokument und Titelaufnahme sind direkt verlinkt.

Die Suchmöglichkeiten über ILTIS, OPAC oder Gateway wurden hinsichtlich der Materialarten erweitert um Online- (mat o) und Offline- (mat s) Publikationen, die Indexierung der URN - einer serverunabhängigen Kennzeichnung elektronischer Dokumente – als neue Suchmöglichkeit ist geplant.

Durch den Link in der Titelaufnahme kann das elektronische Dokument auf [deposit.ddb.de](http://deposit.ddb.de) direkt aufgerufen werden.

#### ↳ Dokumentserver [deposit.ddb.de](http://deposit.ddb.de)

Auf dem Dokumentserver Der Deutschen Bibliothek finden sich nicht nur die gelieferten elektronischen Dokumente, sondern auch Angaben zum Stand der Diskussion um elektronische Publikationen, zu laufenden Projekten zu diesem Komplex usw. Die Dokumente selbst sind derzeit noch nicht über eine Suchmaschine direkt suchbar – dies ist aber vorgesehen.

Für jede Online-Dissertation wird aufgrund der beschreibenden Daten des Ablieferers automatisch eine Frontpage in HTML generiert, die – unsichtbar – auch die mitgelieferten Metadaten nach Dublin Core enthält. Sichtbar sind Titel, Verfasserangabe, Hochschulschriftenvermerk und Abstract; eine Übersicht über Formatvarianten, ggf. notwendige Viewer sowie der Authentisierungs-Code sind dort ebenfalls verankert.

#### ↳ Formatfragen

Hinsichtlich der Formate hat Die Deutsche Bibliothek eine Präferenzliste aufgestellt, nach der die Mitarbeiter bei parallel vorliegenden Formatversionen die Wahl treffen, welches Dokument ‚geholt‘ werden soll.

Formatkonversionen finden derzeit nicht statt.

Die gängigen Dateiformate werden in zwei Gruppen gegliedert:

- Formate mit vergleichsweise positiven Eigenschaften im Bereich der Langzeitarchivierbarkeit (Gruppe A, Archivierung).  
Die Formate der Gruppe A bieten bessere Voraussetzungen zur künftigen Datenmigration als die Formate der Gruppe B.
- Formate mit vergleichsweise positiven Eigenschaften bei der Präsentation (Gruppe B, Präsentation).  
Die Formate der Gruppe B zeichnen sich durch hohe Präsentationsqualität und gute Viewer-Verfügbarkeit aus.

Die Zuordnung der Formate zu diesen beiden Gruppen und ihre Präferenz sieht bei Der Deutschen Bibliothek derzeit wie folgt aus:

A. Archivierung	B. Präsentation
1. SGML (+)	1. PDF (++)
2. HTML (& ggf. Grafik)	2. PS
3. XML (+)	3. TIFF
4. RTF	4. JPEG
5. Text & Layout (.doc, ...)	5. GIF
6. TeX, LaTeX (.tex, .dvi)	
7. ASCII (.txt)	

Auswahlprinzip:

1. Aus Gruppe A wird das Format mit der höchsten Präferenz gewählt.
  2. Sollte kein Format aus Gruppe A verfügbar sein, wird aus Gruppe B dasjenige mit der höchsten Präferenz gewählt.
- (+) Bei den Formaten SGML und XML wird wegen der derzeit noch schlechten Präsentationsmöglichkeiten (Viewer-Verfügbarkeit) **zusätzlich** das Präsentationsformat mit der höchsten Präferenz ausgewählt.
- (++) Wegen der positiven Eigenschaften des PDF-Formates wird dieses grundsätzlich **zusätzlich** zu einem Archivierungsformat ausgewählt.

#### ↳ Authentifizierung

Die Aufgabe des bei Der Deutschen Bibliothek angewandten Authentifizierungs-verfahrens (MD5 = Message Digest Algorithm 5) ist, folgende Sachverhalte feststellen zu können:

- Übereinstimmung von Dokumentversionen (Gewährleistung, daß eine vorliegende Kopie voll identisch ist mit dem ‚Original‘),
- Beziehung zwischen Dokumentbeschreibung (Metadaten) und Dokument (Gewährleistung, daß sich eine Dokumentbeschreibung auf das vorliegende Dokument bzw. die Kopie bezieht),
- Herkunft des Dokuments (Gewährleistung, daß die vorliegende Dokumentkopie von dem Server stammt, der das Original verwaltet).

#### ↳ Internetzugriff

Der Zugriff von außen auf die Online-Dissertationen auf dem Dokumentserver ist sowohl über den OPAC als auch über das Z39.50 Gateway möglich. Auf der Homepage Der Deutschen Bibliothek ist der Unterschied zwischen beiden Zugriffen kurz skizziert.

An dieser Stelle sei darauf hinzuweisen, daß der Außenzugriff bei den Online-Dissertationen unproblematisch ist, da keine besonderen Zugriffsrechte verwaltet werden müssen. Wie der Außenzugriff bei z.B. Verlagspublikationen geregelt wird, ist noch Verhandlungssache zwischen Der Deutschen Bibliothek und den Verlagen.

### **Ablieferung elektronischer Verlagspublikationen**

Einige grundlegende Aspekte wurden in Hearings mit Verlegern, Bibliothekaren, Informationsspezialisten und Regierungsvertretern bereits im Vorfeld definiert und vom Verlegerausschuß des Börsenvereins des Deutschen Buchhandels im Juni 1997 verabschiedet:

- Netzpublikationen sind grundsätzlich über Datennetze abzuliefern, auf Anforderung ggf. auf Datenträgern.
- Netzpublikationen in verschiedenen Formaten sind in dem von der Bibliothek gewünschten Format abzuliefern.
- Werden inhaltlich identische Medienwerke gleichzeitig in physischer Form und als Netzpublikation verbreitet, sind beide Formen abzuliefern.
- Werden inhaltlich identische Netzpublikationen von mehreren Anbietern gleichzeitig verbreitet, sind sie nur einmal abzuliefern.
- Der Bibliothek ist gestattet, eine Kopie der digitalen Publikation für Zwecke der Langzeitarchivierung herzustellen. Dabei ist die Authentizität des Inhalts zu sichern.

Zur Abstimmung der bibliothekarischen und verlegerischen Belange im Hinblick auf die Ablieferung von Netzpublikationen ist im August 1998 eine kleine Arbeitsgruppe gegründet worden, bestehend aus Vertretern der Verlage Buchhändler-Vereinigung, DuMont, Saur, Springer und Wiley-VCH sowie Vertretern unterschiedlicher Bereiche Der Deutschen Bibliothek. In einem ersten Gespräch im Herbst wurden die nachfolgenden Festlegungen getroffen, die das weitere Vorgehen bestimmen:

- Bereitschaft der Verlage zur Kooperation, wenn
  - die Kosten dafür in einem vertretbaren Rahmen bleiben und
  - die 'Schnittstelle' für die Übergabe der Dokumente organisatorisch und technisch klar definiert werden kann.

Die Langzeitarchivierung elektronischer Veröffentlichungen ist bei den Verlagen in dem diskutierten Umfang nicht möglich, daher besteht ein großes Interesse daran, daß Die Deutsche Bibliothek diese Aufgabe zentral in Deutschland übernimmt.

- Arbeitsprogramm:
  1. Definition des 'Status quo', Bestandsaufnahme -->  
Testzugang zu den Servern der beteiligten Verlage,  
Entwurf eines Fragebogens und Verteilung über den AK Elektronisches Publizieren
  2. Definition der Sammelrichtlinien für statische und dynamische elektronische Publikationen
  3. Formen der 'Erwerbung'
  4. Konventionen
  5. Kostenaspekte
  6. Erschließung und Beschreibung (Metadaten der Verlage)
  7. CIP-Verfahren
  8. Dokumentenspeicherung und -archivierung
  9. Externe Nutzung

Die ersten Ergebnisse der Arbeitsgruppe wurden bereits in ‚Dialog mit Bibliotheken‘ veröffentlicht. In Kürze soll außerdem eine eigene Rubrik auf dem Dokumentserver eingerichtet werden, die ständig über neue Entwicklungen berichtet.