

Jörg Berkemeyer, Berthold Weiß

## **Metadaten für Online-Dissertationen**

### **Neues Format und neue Schnittstelle**

Seit dem 1. Juli 1998 werden Online-Dissertationen und -Habilitationen an Der Deutschen Bibliothek gesammelt. Dies geschieht in Zusammenarbeit mit den Universitätsbibliotheken, die diese auf freiwilliger Basis an Die Deutsche Bibliothek melden. Dies geschah bisher ausschließlich über das auf dem Server Der Deutschen Bibliothek angebotene Anmeldeformular. Seit März 1999 werden den Bibliotheken neue Möglichkeiten der Anmeldung der Publikationen und der Abgabe der begleitenden Metadaten angeboten. Dafür war die Entwicklung eines Metadatenformats und einer Metadatenschnittstelle notwendig.

### **Metadatenformat für Dissertationen und Habilitationen**

Voraussetzung für die Nutzung und Weiterverarbeitung von Metadaten sowohl an den Universitätsbibliotheken als auch bei Der Deutschen Bibliothek ist die Festlegung auf einen allgemeingültigen Metadatenstandard. Das Metadatenformat für Online-Hochschulschriften (gültig für Dissertationen und Habilitationen), genannt METADISS<sup>1)</sup>, ist in Gemeinschaftsarbeit von dem Teilprojekt »Metadaten«<sup>2)</sup> des DFG-Projekts »Dissertationen Online«<sup>3)</sup> und Der Deutschen Bibliothek entwickelt und von Der Deutschen Bibliothek mit Bibliotheksverbänden und einzelnen Bibliotheken abgestimmt worden. METADISS basiert auf dem »Dublin Core Metadata Element Set«<sup>4)</sup>. Der Metadatensatz ist an einigen Stellen erweitert worden, um allen Anforderungen zur Beschreibung der Online-Dissertationen und zur Übermittlung und Weiternutzung der

Daten im Geschäftsgang gerecht zu werden. Als Dokument-Format ist HTML 4.0 vereinbart.

Das Metadatenformat unterscheidet zwischen obligatorischen und freiwilligen Datenelementen. Zwingend vorgeschrieben sind administrative Metadaten, wie Angaben zur Abgebenden Stelle (Name, Sitz, Ansprechpartner und Identifikationsnummer der Institution), aber auch deskriptive Metadaten (Art der Hochschulschrift, Autor, Titel, Sprache, Promotionsdaten), die als Zugriffspunkte für eine Suche nach der Online-Dissertation als notwendig angesehen werden. Für die Übergabe der Online-Dissertation an Die Deutsche Bibliothek und die nachfolgende Archivierung sind auch technische Metadaten des Dokuments erforderlich (URL des Dokuments, Dokument-Format und Anzahl der Dateien). Neben den obligatorischen Datenelementen gibt es auch Datenelemente, zu denen Angaben zwar erwünscht, aber nicht vorgeschrieben sind, so zum Beispiel weitere Informationen zum Verfasser (Adresse, Geburtsdatum und -ort), weitere beschreibende Elemente zur Dissertation (weitere Titelangaben, Schlagworte, Abstract, Gutachter, Betreuer/-Doktorvater), mögliche Beziehungen der Online-Dissertation zu anderen Dokumenten oder eventuelle rechtliche Hinweise im Zusammenhang mit der Veröffentlichung der Dissertation.

### **Metadatenschnittstelle für Dissertationen und Habilitationen**

Ziel der Metadatenschnittstelle<sup>5)</sup> ist es, den Wünschen der Universitätsbibliotheken entgegenzukommen und ihnen eine Vielfalt an Ablieferungs- und Übertragungswegen anzubieten, aus denen sie das Verfahren auswählen können, das ihren eigenen Ge-

schäftsgängen am besten entspricht. Ziel ist es insbesondere, eine doppelte Erfassung der Metadaten zu vermeiden. Um die Metadatenchnittstelle nutzen zu können, müssen die Metadaten dem oben beschriebenen Metadatenformat entsprechen.

Die Universitätsbibliotheken können aus vier Ablieferungsverfahren wählen:

- per HTML-Formular,
- per e-mail,
- per e-mail mit Dateianhang, sowie
- per e-mail mit Link zu einem HTML-Dokument.

Im ersten Fall bietet Die Deutsche Bibliothek ein Anmeldeformular<sup>6)</sup> für Hochschulschriften an, welches ausgefüllt über ein CGI-Skript als e-mail mit den eingetragenen Daten an Die Deutsche Bibliothek versandt wird. Die erfolgreiche Übertragung wird durch die Übersendung einer HTML-Seite mit den eingetragenen Metadaten an den Absender bestätigt.

Ein weiterer Übertragungsweg ist die Übermittlung per e-mail - entweder direkt als Text in der e-mail enthalten oder als ASCII-Text im Dateianhang. Im »Betreff« der e-mail werden die Identifizierungsnummer der abliefernden Stelle, der Name des Autors und die Sachgruppe der Deutschen Nationalbibliographie eingetragen. Dieses Verfahren bietet sich dann an, wenn die Metadaten bei der abgebenden Stelle bereits als Datei vorliegen.

Liegen die Metadaten zu der Online-Dissertation bereits als Datensatz im Internet vor, ist es ausreichend, wenn von der abgebenden Universitätsbibliothek per e-mail lediglich ein Verweis auf diese Webseite an Die Deutsche Bibliothek verschickt wird.

Problemfälle werden automatisch bearbeitet und gegebenenfalls an den Absender mit dem entsprechenden Vermerk zurückgesandt

oder individuell an Der Deutschen Bibliothek bearbeitet. Ist der gelieferte Metadatenatz fehlerhaft oder nicht vollständig, wird die abliefernde Stelle automatisch benachrichtigt und kann einen korrigierten Datensatz übersenden.

### **Ausblick**

Das Verfahren hat sich bewährt. Die abliefernden Bibliotheken nutzen die verschiedenen Übertragungswege. Es sind derzeit ca. 500 Online-Dissertationen und -Habilitationen (Stand: 25. August 1999) an Die Deutsche Bibliothek gemeldet worden.

Das derzeitige Metadatenformat kann noch nicht als abgeschlossen angesehen werden. Die zwei folgenden Fragestellungen spielen eine zentrale Rolle und sollten bei einer Erweiterung des Metadatenformats Berücksichtigung finden.

### **Sammlung und Speicherung technischer Metadaten**

Der Wandel im Bereich der elektronischen Publikationen schreitet rasch voran. Formate werden durch neue Versionen ersetzt, neue Hardware ermöglicht den Einsatz komplexerer Multimedia-Elemente, mit neuer Systemsoftware kommen gleichzeitig neue Anwendungen auf den Markt. Insbesondere im Internet-Bereich sind die Lebenszyklen der Software-Produkte sehr kurz und die Auswahl an dynamischen Elementen sehr umfangreich. Verleger und Autoren nutzen diese neuen Funktionen und Eigenschaften, um interessantere Inhalte zu vermitteln und sich im Wettbewerb abzusetzen. Für Bibliotheken ist es schwierig, mit der Entwicklung mithalten und den Benutzern alle Publikationen zur Verfügung zu stellen. Noch schwieriger ist es, eine langfristige Archi-

vierung und Nutzung der Publikationen zu gewährleisten.

Metadaten spielen bei der Langzeitarchivierung und zukünftigen Nutzung der Publikationen eine zentrale Rolle. Strategien der Langzeitarchivierung wie die Migration der Daten und Anwendungen sowie die Emulation von Hard- und Software erfordern in Zukunft detaillierte bibliographische und technische Informationen. Da heute noch nicht exakt voraussagbar ist, welche Informationen in Zukunft für die Entwicklung eines Emulators oder einer Migrationsstrategie erforderlich sind, gilt es, die technische Systemumgebung und das Dokumentformat möglichst umfassend zu erfassen. Technische Metadaten sind vor allem in folgenden Bereichen relevant

- Medientyp - auf welchem Medium ist die Publikation gespeichert und welches System-Format wurde für die Formatierung verwendet;
- Software-Anforderungen zum Anzeigen des Dokuments bzw. Ausführen der Anwendung;
- Hardware-Anforderungen zum Anzeigen und Ausführen des Dokuments bzw. der zugrundeliegenden Software;
- Dokumentformat

Zur Speicherung der technischen Metadaten bietet sich das Dublin Core Element »Format« an. Die Definition der einzelnen Subelemente ist bis jetzt allerdings noch nicht erfolgt. Eine Absprache der Elemente und technischen Anforderungen auf nationaler und internationaler Ebene ist dringend erforderlich. Entwicklungsarbeit in dieser Richtung erfolgt in Projekten wie »NEDLiB«<sup>7)</sup>. Für den praktischen Einsatz und die Übermittlung der technischen Metadaten gilt der Grundsatz: so einfach wie möglich - so umfassend wie nötig. Bibliotheken,

Autoren, Verlegern und Produzenten müssen Instrumente zur Verfügung gestellt werden, die die Erfassung und Übermittlung dieser zusätzlichen technischen Metadaten erleichtern. Die Metadatenchnittstelle für Online-Dissertationen könnte nach der Definition der technischen Metadaten in einem nächsten Schritt um diese Elemente ergänzt werden.

### **Metadaten für Personen**

Im bisherigen Metadatenformat für Online-Hochschulschriften können bereits jetzt im Feld »DC.Creator« mit weiteren entsprechenden »Qualifiern« Informationen zur Person, wie Name, Geburtsort und -datum, angegeben werden, die sich nicht auf die Dissertation, sondern speziell auf die jeweilige Person beziehen. Der Gedanke liegt nahe, diese Angaben analog der Differenzierung von Norm- und Titeldaten in bibliographischen Formaten in einem getrennten Metadatenatz anzugeben, der die Person beschreibt.

Die Deutsche Bibliothek plant im Rahmen des von der Deutschen Forschungsgemeinschaft geförderten Projekts META-LIB<sup>8)</sup> die Entwicklung eines solchen Metadatenatzes für Personen. Der Verfasser der Publikation wird aufgefordert, neben den Metadaten zu seiner Publikation, auch Metadaten zu seiner Person - auf freiwilliger Basis - mitzuliefern. Diese Metadaten könnten dann in entsprechender Auswahl auch für die Personennamendatei <PND> weitergenutzt und im Internet zur Recherche angeboten werden. Als Grundlage für die Datenelemente im Personenmetadatenatz soll das »Dublin Core Metadata Element Set« dienen. Da die Dublin-Core-Gemeinschaft für diesen Bereich noch keinen Elementsatz entwickelt hat, ist beabsichtigt,

eigene Ideen als Vorschläge in die »Dublin Core Working Groups« zu bringen und mit anderen Entwicklungen abzustimmen. Bereits in diesem Bereich existierende Anwendungen werden geprüft und in die eigene Entwicklungsarbeit einbezogen werden. Beispiele für existierende Anwendungen sind zum Beispiel das Math-Net-Projekt<sup>9)</sup>, in dem ein DC-Metadatenatz für »Personal Homepages« erstellt worden ist<sup>10)</sup>, oder die »vCard«, eine elektronische Visitenkarte, die vor allem im Bereich von E-Mail-Nachrichten angewendet wird<sup>11)</sup>.

---

1)

<<http://deposit.ddb.de/metadiss.htm>>

2)

<<http://www.ub.uni.duisburg.de/dissonline/>>

3)

<[http://www.educat.hu-berlin.de/diss\\_online/index.html](http://www.educat.hu-berlin.de/diss_online/index.html)>

4)

<<http://purl.org/DC/>>

5)

Eine detaillierte Beschreibung der Metadatenchnittstelle finden Sie im WWW unter  
<<http://deposit.ddb.de/index.htm>>

6)

<<http://www.ddb.de/service/anmeldeformulare/hochschulschrift.htm>>

7)

Networked European Deposit Library  
<<http://www.konbib.nl/nedlib/>>

8)

Metadaten-Projekt deutscher Bibliotheken, Teilprojekt Der Deutschen Bibliothek  
<<http://www.ddb.de/partner/metalib.htm>>

9)

<<http://www.math-net.de/>>

10)

<<http://www.math-net.de/project/metadata/personal.html>>

11)

<<ftp://ftp.isi.edu/in-notes/rfc2426.txt>>