

3. META-LIB Workshop an der SUB Göttingen

Metadata: New developments - new frontiers

Hans J. Becker, Petra Lepschy, Inka Tappenbeck

Am 22. und 23. Mai 2000 fand an der Niedersächsischen Staats- und Universitätsbibliothek Göttingen unter internationaler Beteiligung der 3. META-LIB-Workshop statt. Unter dem Titel „*Metadata: New developments - new frontiers*“ <<http://www2.sub.uni-goettingen.de/metalib/index.html>> diskutierten Vertreter von Hochschulen, Bibliotheken, Museen, Rechenzentren und anderen Kultur- und Informationseinrichtungen gemeinsam über aktuelle Fragen und Entwicklungen im Metadatenbereich.

Diese Veranstaltung schließt an die ersten beiden META-LIB-Workshops an, die 1998 in der Niedersächsischen Staats- und Universitätsbibliothek Göttingen und 1999 an der Deutschen Bibliothek in Frankfurt stattfanden. Während sich die Diskussion vor zwei Jahren um „*Neue Perspektiven für die Erschließung von Netzpublikationen in Bibliotheken*“ drehte und im vergangenen Jahr dann „*Bibliographische Metadaten im Spannungsfeld von Heterogenität und Standardisierung*“ im Zentrum des Interesses standen, konzentrierte sich der dritte Workshop auf neue Entwicklungen und Probleme bei der Anwendung von Metadaten im Kontext von Bibliotheken, Museen und Archiven. Wie auch in den vergangenen Jahren, stellte dabei die Frage der Anwendbarkeit des Dublin Core Metadata Element Set <<http://purl.org/DC/documents/rec-dces-19990702.htm>> für diese verschiedenen Aufgabenbereiche einen wichtigen Schwerpunkt dar. In diesem Zusammenhang wurde auch die Frage der Notwendigkeit von Standardisierungen für die Anwendung des Dublin Core Set sowie die Möglichkeit der Interoperabilität mit anderen bibliographischen Beschreibungsstandards diskutiert. Auch Vorträge, in denen Beispiele für Modifikationen und Erweiterungen des Dublin Core Set für spezifische Anwendungsbedingungen dargestellt wurden, trugen zur Vielfalt der auf diesem Workshop vorgestellten Themen bei.

Den Auftakt des Workshops bildete die Begrüßung der Teilnehmer durch *Prof. Dr. Elmar Mittler*, den Leitenden Direktor der Niedersächsischen Staats- und Universitätsbibliothek Göttingen. Mittler betonte die Bedeutung der Metadatenentwicklung für die Zukunft der globalen wissenschaftlichen Informationsversorgung sowie für die Archivierung unseres kulturellen und wissenschaftlichen Erbes. Die sich daran anschließenden Vorträge, Diskussionen und Arbeitsgruppentreffen befassten sich vor allem mit folgenden Themenschwerpunkten:

1. Dublin Core für Bibliotheken, Museen und Archive
2. Metadatenstandardisierung und Registry
3. Metadaten, Regelwerke, Datenformate

Diese Themengebiete wurden, erweitert um den Komplex „RDF und XML in Bibliotheken“, in vier zusätzlichen Arbeitsgruppentreffen vertieft, deren Ergebnisse anschließend im Plenum vorgestellt wurden.

Der erste Themenbereich, in dem es um die Anwendung des Dublin Core Metadata Element Set in Bibliotheken, Museen und Archiven ging, wurde durch den Vortrag von *Angela Spinasse* vom Consortium for the Computer Interchange of Museum Information (CIMI) <<http://cimi.org>> eingeleitet. Spinasse gab einen einführenden Überblick über die Anwendung von Metadaten im Museumsbereich, erläuterte die spezifischen Probleme, die sich dabei besonders aufgrund der Heterogenität der dort vorhandenen Ressourcen erge-

ben und skizzierte die besonderen Anforderungen, die die Metadatenbeschreibungen für den Zugriff auf diese Bestände erfüllen müssen. Sie stellte die CIMI Dublin Core Testdatenbank vor, die ca. 200.000 Dublin Core-Datenätze im XML/DTD-Format enthält und erklärte deren Funktionen für das Retrieval und den Datenaustausch mit anderen Einrichtungen.

An diese Präsentation schloss sich der Vortrag von *Jutta Weber* von der Staatsbibliothek zu Berlin - Preußischer Kulturbesitz an. Weber stellte das EU-Projekt MALVINE (Manuscripts and Letters via Integrated Networks in Europe) <<http://www.malvine.org>> vor, in dem es um die Entwicklung einer Suchmaschine zu Nachlass- und Autographenbeschreibungen und eines gemeinsamen Metadatenformats für Nachlässe und Autographen geht. Auch in diesem Projekt liegt das Dublin Core Metadata Element Set zugrunde, es ist jedoch an einigen Stellen an die spezifischen Erfordernisse der Beschreibung dieser Dokumentarten angepasst. So steht beispielsweise an Stelle des „DC.Creator“ und „DC.Contributor“ in MALVINE der „agent“ als übergeordnete Beschreibungskategorie, die auch in der Lage ist, die im Dublin Core nicht zu erfassenden Adressaten von Briefen aufzunehmen und damit suchbar zu machen. Weiterhin stellte Weber die verschiedenen in MALVINE möglichen Suchstrategien vor und wies auf die bestehenden Desiderata in der Verwendung von Metadaten für die Beschreibung von Nachlässen und Autographen, wie etwa die umfassende Nutzung von Normdaten, hin.

Den dritten Vortrag zu diesem Themenbereich hielt *Michael Day*, Research Officer an dem an der University of Bath angesiedelten UK Office for Library and Information Networking (UKOLN) <<http://www.ukoln.ac.uk>>. Day stellte das Projekt Cedars (CURL exemplars in digital archives) <<http://www.leeds.ac.uk/cedars/index.htm>> vor, in dem es um die Entwicklung eines Modells zur Archivierung digitaler Ressourcen geht. Nach einer Übersicht über die wesentlichen technischen, rechtlichen und organisatorischen Probleme bei der Archivierung digitaler Ressourcen erläuterte er das von Cedars in Anlehnung an das OAIS-Modell (Reference Model for an Open Archival Information System) <<http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-1.pdf>> entwickelte Metadatenmodell, das alle für die Archivierung relevanten Informationen beinhalten soll <<http://www.leeds.ac.uk/cedars/MD-STR-5.pdf>>. Des Weiteren stellte Day das OAIS-Prozessmodell vor <<http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-1.pdf>>, in dem die einzelnen Archivierungsfunktionen beschrieben und bestimmten Arbeitsprozessen zugeordnet sind.

Den zweiten Themenkomplex „Metadatenstandardisierung und Registry“ leitete *Rachel Heery*, Assistant Director (Research and Development) des UKOLN, mit ihrem Vortrag über „Dublin Core and registries“ ein. Heery verdeutlichte die Notwendigkeit von Metadaten Registries angesichts der Viel-

zahl und Heterogenität der im Internet verbreiteten Metadatenformate und stellte die im Projekt DESIRE (Development of a European Service for Information on Research and Education) <<http://www.desire.org>> entwickelte Metadaten Registry Datenbank vor. Wie auch die Datenbank „Meta-Form“ der Niedersächsischen Staats- und Universitätsbibliothek <<http://www2.sub.uni-goettingen.de/metaform/index.html>>, will diese Registry Datenbank über vorhandene Formate informieren und die Möglichkeit bieten, sie über crosswalks zu vergleichen, um den Gebrauch einheitlicher Standards zu befördern und die Interoperabilität zwischen den verwendeten Metadatenformaten zu erhöhen.

Dieser Beitrag wurde inhaltlich ergänzt und erweitert durch den Vortrag *Maks Dekkers*' von PricewaterhouseCoopers <<http://www.pricewaterhousecoopers.org/gx/eng/main/home/index.html>> über „Metadata, Dublin Core and standardisation“. Auch Dekkers betonte die Notwendigkeit einer Koordination im Bereich der Metadatenentwicklung und -verwendung durch Registries, mit dem Ziel, eine weltweite, bereichsübergreifende Interoperabilität der im Internet verwendeten Metadatenformate herzustellen. Er gab einen Überblick über die verschiedenen aktuellen Standardisierungsbestrebungen zum Gebrauch von Metadaten, wie zum Beispiel im europäischen Bereich der CEN/ISSS Workshop <<http://www.cenorm.be/iss/Workshop/MMI-DC>> und auf internationaler Ebene der NISO Draft Standard Z39.85-200X: The Dublin Core Metadata Element Set <<http://www.niso.org/Z3985.html>>. Abschließend forderte Dekkers die Metadaten verwendenden Institutionen auf, lokal zu denken, aber global zu handeln, und die Konvergenz und Interoperabilität ihrer Datenformate stets im Blick zu behalten.

Der dritte Themenbereich „Metadaten, Regelwerke, Datenformate“ wurde von *Eeva Murtomaa* von der Helsinki University Library <<http://www.lib.helsinki.fi>> eingeleitet. In ihrem Vortrag „Dublin Core and Functional Requirements for Bibliographic Records“ stellte Murtomaa das von der IFLA Study Group on the Functional Requirements for Bibliographic Records entwickelte Modell, das sich besonders durch seinen objektorientierten Ansatz auszeichnet, vor <<http://ifla.org/VIII/s13/frbr/frbr.pdf>>. Danach ist jedes Objekt auf bis zu vier Ebenen beschreibbar: als „work“ (Werk), „expression“ (intellektuelle oder künstlerische Ausdrucksform), „manifestation“ (physische Verkörperung) und „item“ (Exemplar). So ist zum Beispiel das Werk „Gone with the wind“ als originäre geistige Schöpfung von Margaret Mitchell nach dem IFLA-Modell als „work“ zu charakterisieren, seine Umsetzung in Form eines Theaterstücks oder Films als dessen „expression“, die wiederum ihre „manifestation“ in Gestalt einer besonderen Inszenierung oder filmischen Realisierung findet, dessen „item“ dann das einzelne Exemplar, die Filmrolle bzw. die einzelne Auf-

führung ist. Eine weitere Gruppe von Entitäten, beschrieben als „person“ (Person) und „corporate body“ (Institution), wird in diesem Modell zur Darstellung der Beziehungszusammenhänge verwendet, in denen ein Werk stehen kann. Eine dritte Gruppe umfasst mit „concept“, „object“, „event“ und „place“ die Entitäten, über die sich der Inhalt eines Werkes näher bestimmen lässt. Die möglichen Beziehungen dieser Entitäten zu- und ihre Abhängigkeiten voneinander werden in dem sogenannten „entity-relationship-model“ anschaulich abgebildet. Diese Initiative der IFLA zielt darauf ab, ein solches Beziehungsmodell in Informationsräumen zu implementieren und den Benutzern damit eine komfortable Möglichkeit der Suche nach gewünschten Ressourcen an die Hand zu geben. Murtomaa verließ der Hoffnung Ausdruck, dass dieses Modell zukünftig nicht nur in Bibliotheken, sondern auch in Verlagen, Museen und Archiven Anwendung finden und damit eine gemeinsame Datennutzung über die Grenzen der Bibliothekswelt hinaus möglich sein wird.

Rebecca Guenther vom Network Development and MARC Standards Office der Library of Congress <<http://lcweb.loc.gov/marc/ndmso.html>> stellte in ihrem Vortrag „Dublin Core and MARC 21“ zunächst die Funktionalitäten dieser beiden Formate als Beschreibungsinstrumente für elektronische Ressourcen dar. Daran anschließend erläuterte sie die Möglichkeiten des mappings, also der Abbildung eines Feldersystems auf ein anderes, zwischen MARC und Dublin Core. Guenther hob die durch mappings zu erzielenden Synergieeffekte hervor, die vor allem durch die Möglichkeit, in strukturell verschiedenen Formaten mit denselben Daten zu arbeiten, entstehen. Sie wies aber auch auf die Grenzen dieses Instruments hin: Der Grad der durch mappings zu erzielenden Datenkonvergenz ist unmittelbar von der Bedeutungsähnlichkeit der Felder der zugrundeliegenden Datenformate abhängig. Elemente eines Datenformats, deren Bedeutungsgehalte in einem anderen Format nicht als Beschreibungseinheiten ausgewiesen sind, lassen sich darin auch nicht abbilden. Dies gilt insbesondere für mappings zwischen komplexen und einfach strukturierten Formaten sowie zwischen solchen, deren Semantik auf die Beschreibung unterschiedlicher Arten von Ressourcen zugeschnitten ist. Guenther betonte daher die Bedeutung von vereinheitlichenden Standards in der Anwendung von Datenformaten und ebenso von crosswalks, also Schemata, die die semantische Konvergenz zwischen verschiedenen Formaten abbilden <<http://www2.sub.uni-goettingen.de/metaform/crosswalks.html>>. In diesem Zusammenhang hob sie die Funktion des Dublin Core Metadata Element Set als „kleinster gemeinsamer Nenner“ verschiedener Metadatenformate hervor, über den eine Nutzung der zentralen Beschreibungsdaten auch zwischen Anbietern heterogener Daten möglich sei. Um die Interoperabilität in diesem Bereich in Zukunft weiter zu erhöhen, forderte Guenther eine stärkere theoretische und praktische Auseinandersetzung der Bibliotheken mit den

Möglichkeiten, die mit der Verwendung des Dublin Core Metadata Element Set verbunden sind.

Die Reihe der Vorträge wurde mit einem Beitrag von *Hans J. Becker*, Niedersächsische Staats- und Universitätsbibliothek Göttingen <http://www.sub.uni-goettingen.de/java_home.htm> beschlossen, der in Vertretung der erkrankten Referentin *Diann Rusch-Feja* von der Bibliothek des MPI für Bildungsforschung <<http://www.mpib-berlin.mpg.de/>> die Arbeit der DC-Working Group Education <<http://purl.org/DC/groups/education.htm>> und die Genese des erweiterten Dublin Core Set zu Education-Materialien vorstellte. Diese Entwicklung, die im Herbst 1999 kurz vor dem DC-7 Meeting in der DDB in Frankfurt <<http://www.ddb.de/>> begann, wurde als prototypisches Beispiel für die Mechanismen der internationalen Entwicklung von speziellen Dublin Core Anwendungen mit Hilfe einzelner Entwicklungsschritte charakterisiert. In diesem Kontext betonte Becker, dass sich an diesem Prozess abermals eine der großen Stärken der DC-Entwicklung und Community erwiesen habe: Innerhalb kurzer Zeit gelang es einer Reihe von Institutionen und Anwendergruppen, sich auf ein gemeinsames, um spezifische Kategorien erweitertes Dublin Core Format zu einigen, wodurch sowohl Interoperabilität als auch Crosssearching in internationalen Datenbanken ermöglicht wird.

Die Inhalte der auf diesem Workshop gehaltenen Vorträge wurden in vier an verschiedenen thematischen Schwerpunkten orientierten Arbeitsgruppen diskutiert und vertieft. Besonders hervorzuheben ist hier die Arbeit der Gruppe „registries“, in der ein internationaler Konsens bezüglich der Vorgehensweise und der Terminologie erzielt und dokumentiert wurde. Mit einer kurzen Darstellung des Stands des DFG-Projekts META-LIB durch die beiden Partner Die Deutsche Bibliothek, Frankfurt, und die Niedersächsische Staats- und Universitätsbibliothek Göttingen fand der 3. META-LIB Workshop seinen Abschluss.

