Deutsche Initiative für
Netzwerkinformation (DINI)

Elektronisches Publizieren an Hochschulen

Inhaltliche Gestaltung der OAI-Schnittstelle

- Empfehlungen -

Arbeitsgruppe Open Archives Initiative in Deutschland

Oktober 2003

Inhaltsverzeichnis

	Zusammenfassung	3
1	Das Open Archives Protocol for Metadata Harvesting (OAI-PMH)	3
2	Empfehlungen zur Definition von Sets	5
2.1	Allgemeine Bemerkungen	5
2.2	Inhaltliche Gliederung	6
2.3	Gliederung nach dem Publikationstyp	9
2.4	Gliederung nach dem Dokumenttyp	10
3	Weitere Empfehlungen	10
3.1	Dublin Core und andere Metadatenformate	10
	Impressum	12

Zusammenfassung

Das *Open Archives Protocol for Metadata Harvesting* (OAI-PMH) erlaubt es, eigene Metadaten, die der Beschreibung beliebiger Objekte dienen, mit anderen zu teilen. Dieses Papier enthält neben einer kurzen Übersicht über das Protokoll Empfehlungen für die Verwendung von Sets (Mengen) durch deutsche Datenprovider sowie für die Verwendung der Metadatenfelder von *Dublin Core* (DC). Dabei steht das Ziel im Mittelpunkt, einen effizienten Austausch von Metadaten zwischen den Nutzern des OAI-Protokolls zu gewährleisten.

1 Das Open Archives Protocol for Metadata Harvesting (OAI-PMH)

Das *Open Archives Protocol for Metadata Harvesting* (OAI-PMH) ermöglicht einen effizienten Austausch von Metadaten und impliziert eine funktionale Aufteilung in Anbieter von (Dokumenten und) Metadaten, so genannte *Datenprovider*[1], und darauf aufbauende Dienste (*Serviceprovider*[2]). Das OAI-PMH basiert auf dem grundsätzlichen Prinzip des so genannten *Harvesting*[3], bei dem im Gegensatz zum Ansatz des *Cross Searching*[4] eine asynchrone Suche durchgeführt wird. Das heißt, der Serviceprovider fragt in regelmäßigen Abständen die Metadaten der Dataprovider ab und speichert diese in seiner lokalen Datenbank. Konkrete Suchanfragen (z.B. von Endnutzern) werden im Falle des Harvesting-Ansatzes ausschließlich mithilfe der Datenbank beantwortet.

Das OAI-Protokoll basiert auf weithin bekannten und verbreiteten Standards. Es setzt auf das *Hypertext Transfer Protocol* (HTTP) auf und verwendet zur Kodierung der Metadaten und der sonstigen in den Antworten enthaltenen Informationen die *eXtensible Markup Language* (XML). Wiewohl sich das OAI-Protokoll zur Übertragung von Metadaten in beliebigen (durch ein XML-Schema definierten) Formaten eignet, ist *Dublin Core* aus Gründen der Interoperabilität als Minimalstandard in das OAI-PMH aufgenommen worden. OAI-kompatible Datenprovider müssen in der Lage sein, für ihre Metadaten zumindest Dublin Core auszuliefern. Damit sind die Kommunikation und der tatsächliche Austausch von Metadaten zwischen beliebigen OAI-kompatiblen Daten- und Serviceprovidern ohne weitere zusätzliche Vereinbarungen sofort möglich.

Die im OAI-PMH definierte prinzipielle Trennung zwischen Datenprovider und Serviceprovider schließt natürlich die Entwicklung von Diensten nicht aus, die beide Funktionalitäten beinhalten. Diese Möglichkeit wird von so genannten *aggregierenden* Datenprovidern ausgenutzt. Sie fragen über das OAI-Protokoll die verfügbaren Daten einer bestimmten Menge von Datenprovidern ab und halten diese für Anfragen anderer Serviceprovider ebenfalls über eine OAI-Schnittstelle vor.

Für Benutzerinnen und Benutzer von auf dem OAI-PMH basierenden Diensten ist die Technologie, auf der die Suchanfragen beruhen, in der Regel transparent. Ihnen steht zum Beispiel eine Web-Schnittstelle zur Verfügung, über die sie mit dem Serviceprovider kommunizieren und dessen Dienste nutzen. Dass sich die letztlich gefundenen digitalen Objekte auf verteilten Servern befinden, wird aus Nutzersicht erst bei deren Abruf bzw. einer Autorisierungsanforderung sichtbar (siehe Abbildung 1).

---

[1] Als Datenprovider wird eine OAI-kompatible Schnittstelle zu einer Datenbank verstanden, in der sich Metadaten über Dokumente oder andere digitale Objekte befinden und die über eine HTTP-Verbindung erreichbar ist. Sie muss dazu in der Lage sein, OAI-Anfragen entsprechend der Protokolldefinition korrekt zu beantworten.

[2] Der Serviceprovider bietet mithilfe von Daten, die er unter Nutzung des OAI-PMH gesammelt hat, einen (innerhalb der Protokollspezifikation nicht näher definierten) Dienst an. Der aus Sicht des Protokolls relevante Teil des Serviceproviders, der so genannte *Harvester*, versendet OAI-konforme Anfragen an Datenprovider und wertet die entsprechenden Antworten aus.

[3] engl. ernten. Dabei werden regelmäßig alle verfügbaren bzw. relevanten Daten aus den verwendeten Datenquellen abgefragt (geerntet) und in einer zentralen Datenbank gespeichert, innerhalb derer dann die eigentliche Suche erfolgt.

[4] bekannt auch unter dem Namen *Federated Searching*, bezeichnet die unmittelbare Suche in allen verwendeten Datenquellen, wird beispielsweise häufig von Meta-Suchmaschinen verwendet

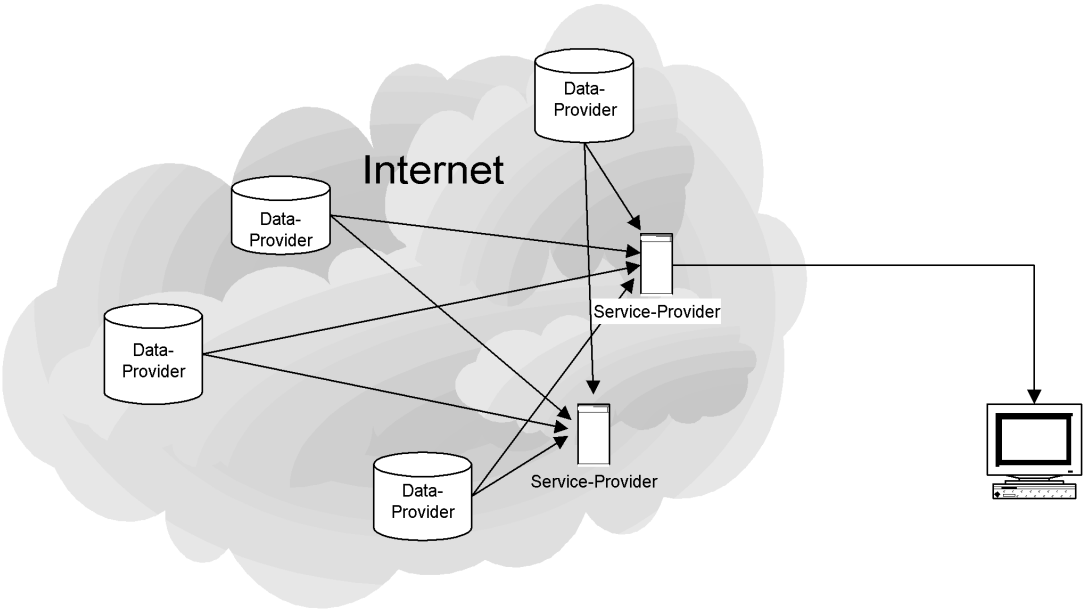

Abbildung 1: Zusammenwirken von Daten- und Serviceprovidern

Das OAI-Protokoll ist kein Suchprotokoll. Qualifizierte Suchanfragen lassen sich über das Protokoll demnach nicht formulieren. Die eigentliche dem Endnutzer oder einem Suchprotokoll zur Verfügung gestellte Suche ist Teil des Serviceproviders und bezieht sich immer auf dessen Datenbank. Die Möglichkeiten, Selektionskriterien bei OAI-Protokollanfragen zu nutzen, beschränken sich auf das letzte Änderungsdatum der Metadaten (**from**- und **until**-Argument) und eine grobe logische Gliederung der Datenbestände in unterschiedliche Mengen (**set**-Argument[5]), deren Definition allerdings nicht in der Spezifikation des OAI-PMH enthalten ist, sondern den jeweiligen Datenprovidern überlassen ist.

Diese Eigenschaft des OAI-PMH ermöglicht es, Metadaten von Datenprovidern selektiv abzufragen. Beispielsweise lassen sich auf diese Weise Fachportale effizient realisieren, da der entsprechende Serviceprovider schon auf Protokollebene die Menge der angeforderten Daten grob einschränken kann. Abbildung 2 zeigt schematisch das Zusammenwirken von Daten- und Serviceprovidern auf der Grundlage einheitlicher Definitionen über die logische Struktur, die so genannte Set-Hierarchie. Jeder Serviceprovider fordert nur die für ihn interessanten Daten an.

Um ein hohes Maß an Interoperabilität innerhalb der dokumentarischen und bibliothekarischen Anwendungen zu gewährleisten und damit auch den Aufbau von Daten- und Serviceprovider-strukturen zu erleichtern, erscheint es sinnvoll, für die Definition und Verwendung von Set-Hierarchien Empfehlungen und Richtlinien zu entwickeln und für die Benutzung dieser Empfehlung zu werben. Dadurch wird es möglich, dass Serviceprovider gezielt Daten nach bestimmten formalen (z.B. Dissertationen) und inhaltlichen Kriterien (z.B. Physik) sammeln und spezifische Dienste (z.B. ein Nachweissystem für Dissertationen oder eine Suchmaschine für Physik-Dokumente) aufbauen können (siehe Abbildung 2).

Die DINI-Arbeitsgruppe *Elektronisches Publizieren* empfiehlt neben der Benutzung des OAI-Protokolls selbst eine inhaltliche und formale Strukturierung des Archivs, um den Aufbau spezifischer Dienste auf der Basis des OAI-PMH zu erleichtern. Diese Empfehlungen bilden den Schwerpunkt des vorliegenden Papiers und werden im folgenden Kapitel ausführlich dargelegt.

[5] set, engl. = Menge

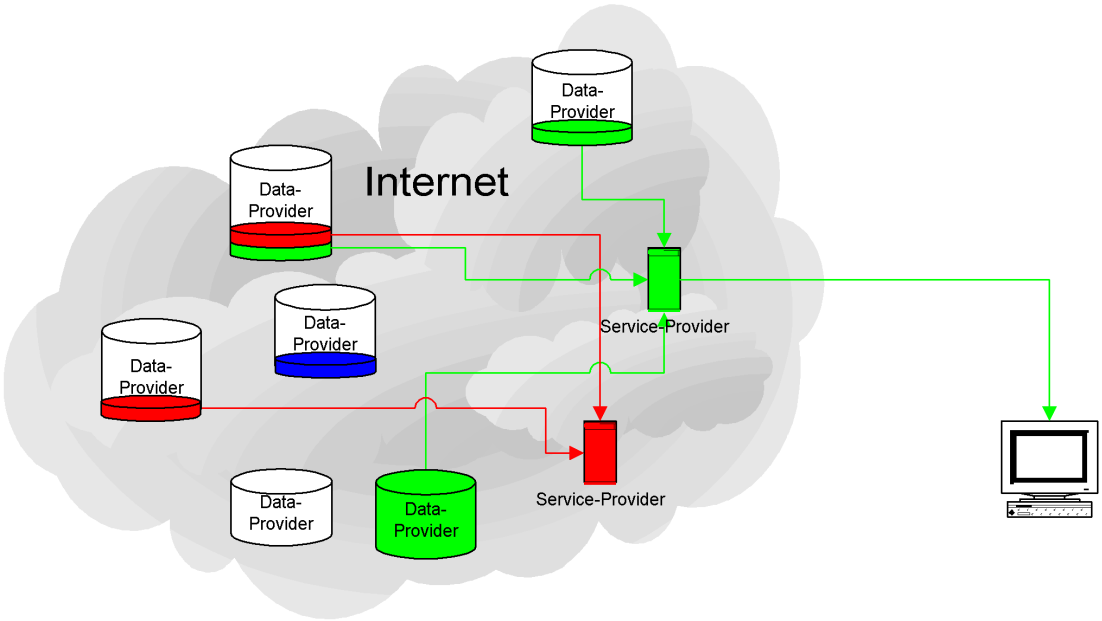

Abbildung 2: Zusammenwirken von inhaltlich oder formal strukturierten Daten- und Serviceprovidern

2 Empfehlungen zur Definition von Sets

2.1 Allgemeine Bemerkungen

Ausgehend von einer Vielzahl möglicher auf dem OAI-Protokoll basierender Dienste ist eine Strukturierung der Archive sowohl nach formalen als auch nach inhaltlichen Kriterien sinnvoll. Die inhaltliche Beschreibung dient einer groben fachlichen Zuordnung der mit den Metadaten beschriebenen Dokumente und Objekte und orientiert sich an den Sachgruppen der Deutschen Bibliothek[6]. Die formalen Untergliederungen beziehen sich auf die Publikationsform und den technischen Dokumenttyp des jeweiligen Objekts.

Die durch einen Datenprovider zur Verfügung gestellten Sets sind maschinell abfragbar. Eine entsprechende OAI-Anfrage liefert neben einem eindeutigen Bezeichner, der in dem **setSpec**-Element enthalten ist, eine verbale Beschreibung (**setName**-Element) des jeweiligen Sets. Damit eine Nutzung des Sets unabhängig von der Kenntnis der deutschen Sprache gewährleistet ist, sollte die Beschreibung vornehmlich in englischer Sprache erfolgen.

Die hier empfohlenen Sets entsprechen drei unterschiedlichen Gliederungsansätzen – einer inhaltlichen Gliederung, einer Gliederung gemäß der Publikationsform und einer Gliederung nach Dokumenttypen. Sie sind jeweils als zweistufige Hierarchie definiert, wobei die Hierarchieebenen gemäß der Spezifikation des OAI-PMH durch einen Doppelpunkt getrennt werden. Ingesamt werden damit drei Sets der obersten Hierarchieebene empfohlen, und zwar die Sets **ddc**, **pub-type** und **doc-type**. Die Sets der zweiten Hierarchieebene (z.B. **ddc:004**) sind als Teilmengen der jeweiligen Sets der obersten Ebene zu verstehen. Abgesehen von diesem hierarchischen Zusammenhang ist mit der Zugehörigkeit eines Objekts zu einem bestimmten Set über die Zugehörigkeit desselben Objekts zu einem anderen Set nichts gesagt. Wünschenswert im Sinne dieser Empfehlungen und logisch durchaus nachvollziehbar wäre es, wenn jeder Metadatensatz in mindestens je einem Set der zweiten Hierarchieebene jedes Gliederungsansatzes vertreten wäre – also beispielsweise eine Dissertation auf dem Fachgebiet der Medizin in den Sets **ddc:610**, **pub-type:dissertation** und **doc-type:text**. Denkbar ist es darüber hinaus auch, dass ein Metadatensatz in mehreren Sets desselben Gliederungsansatzes enthalten ist, also beispielsweise in zwei Sets der inhaltlichen Gliederung

[6] Die Deutsche Bibliothek stellt die Sachgruppen-Gliederung zum Bibliografie-Jahrgang 2004 auf ein auf der Dewey-Dezimalklassifikation (DDC) beruhendes System um. Die bisher geltende Deutsche Nationalbibliografie (DNB) wird nicht mehr verwendet.

(z.B. **ddc:004** und **ddc:610** bei einer Arbeit, die sich mit Informationstechnik in der Medizin beschäftigt). In der Regel wird es darüber hinaus auch Fälle geben, in denen ein Metadatensatz nicht bezüglich jeder Kategorie in einem Set vertreten ist. Die Gründe hierfür können in einer unvollständigen Metadatenerfassung oder in der Migration älterer Metadaten liegen.

Generell wird empfohlen, auf die **ListSets**-Anfrage, die nach allen verfügbaren Sets eines Datenproviders fragt, nur diejenigen Sets zurückzugeben, in denen auch mindestens ein Metadatensatz enthalten ist. Ein Datenprovider, der die vorliegenden Empfehlungen umsetzt, wird also in der Regel auf eine solche Anfrage nicht eine komplette Liste der in den folgenden Abschnitten aufgeführten Sets zurückgeben, sondern nur die aktuell verwendeten Sets liefern.

2.2 Inhaltliche Gliederung

Als inhaltliche Gliederung OAI-kompatibler Datenprovider werden die von der Deutschen Bibliothek verwendeten Sachgruppen gemäß der Dewey-Dezimalklassifikation (DDC) empfohlen. Diese erlauben eine grobe fachliche Einordnung und bieten eine entsprechende Selektionsmöglichkeit für fachlich ausgerichtete Serviceprovider. Tabelle 1 zeigt die gemäß dieser Gliederung definierten Sets. Die erste Spalte beinhaltet den jeweiligen Bezeichner (**setSpec**), in der zweiten Spalte findet sich die Beschreibung des Sets, die auf die OAI-Anfrage ListSets hin im **setName**-Element enthalten ist. Die letzte Spalte ist hier nur der Klarheit halber eingefügt worden. Deren Inhalt spielt für die Protokollanfragen und –antworten keine Rolle.

Die erste Zeile von Tabelle 1 enthält das Set der obersten Hierarchieebene. In ihm enthaltene Metadatensätze beschreiben Dokumente, die eine DDC-konforme Klassifikation besitzen. Alle anderen Zeilen enthalten Sets der zweiten Hierarchieebene, die jeweils Teilmengen des Sets **ddc** bilden.

Tabelle 1: Bezeichnung und Beschreibung der Sets gemäß der inhaltlichen Gliederung

setSpec	setName	Deutschsprachige Beschreibung
ddc	DDC classified objects	Gemäß DDC klassifizierte Objekte
ddc:000	Generalities, Science	Allgemeines, Wissenschaft
ddc:004	Data processing Computer science	Informatik
ddc:010	Bibliography	Bibliografien
ddc:020	Library & information sciences	Bibliotheks- und Informationswissenschaft
ddc:030	General encyclopedic works	Enzyklopädien
ddc:050	General serials & their indexes	Zeitschriften, fortlaufende Sammelwerke
ddc:060	General organization & museology	Organisationen, Museumswissenschaft
ddc:070	News media, journalism, publishing	Geografie, Reisen
ddc:080	General collections	Allgemeine Sammelwerke
ddc:090	Manuscripts & rare books	Handschriften, seltene Bücher
ddc:100	Philosophy	Philosophie
ddc:130	Paranormal phenomena	Parapsychologie, Okkultismus
ddc:150	Psychology	Psychologie
ddc:200	Religion	Religion, Religionsphilosophie
ddc:220	Bible	Bibel
ddc:230	Christian theology	Theologie, Christentum
ddc:290	Other & comparative religions	Andere Religionen
ddc:300	Social sciences	Sozialwissenschaften, Soziologie
ddc:310	General statistics	Statistik
ddc:320	Political science	Politik
ddc:330	Economics	Wirtschaft

setSpec	setName	Deutschsprachige Beschreibung
ddc:340	Law	Recht
ddc:350	Public administration	Öffentliche Verwaltung
ddc:355	Military science	Militär
ddc:360	Social services; association	Soziale Probleme, Sozialarbeit
ddc:370	Education	Erziehung, Schul- und Bildungswesen
ddc:380	Commerce, communications, transport	Handel, Kommunikation, Verkehr
ddc:390	Customs, etiquette, folklore	Ethnologie
ddc:400	Language, Linguistics	Sprachwissenschaft, Linguistik
ddc:420	English	Englisch
ddc:430	Germanic	Deutsch
ddc:439	Other Germanic languages	Andere germanische Sprachen
ddc:440	Romance languages French	Französisch, romanische Sprachen allgemein
ddc:450	Italian, Romanian, Rhaeto-Romantic	Italienisch, Rumänisch, Rätoromanisch
ddc:460	Spanish & Portugese languages	Spanisch, Portugiesisch
ddc:470	Italic Latin	Latein
ddc:480	Hellenic languages Classical Greek	Griechisch
ddc:490	Other languages	Andere Sprachen
ddc:500	Natural sciences & mathematics	Naturwissenschaften
ddc:510	Mathematics	Mathematik
ddc:520	Astronomy & allied sciences	Astronomie
ddc:530	Physics	Physik
ddc:540	Chemistry & allied sciences	Chemie
ddc:550	Earth sciences	Geowissenschaften
ddc:560	Paleontology Paleozoology	Paläontologie
ddc:570	Life sciences	Biowissenschaften, Biologie
ddc:580	Botanical sciences	Pflanzen (Botanik)
ddc:590	Zoological sciences	Tiere (Zoologie)
ddc:600	Technology (Applied sciences)	Technik
ddc:610	Medical sciences Medicine	Medizin
ddc:620	Engineering & allied operations	Ingenieurwissenschaften
ddc:630	Agriculture	Landwirtschaft, Veterinärmedizin
ddc:640	Home economics & family living	Hauswirtschaft
ddc:650	Management & auxiliary services	Management
ddc:660	Chemical engineering	Technische Chemie
ddc:670	Manufacturing	Industrielle Fertigung
ddc:690	Buildings	Hausbau, Bauhandwerk
ddc:700	The arts	Künste, Bildende Kunst allgemein
ddc:710	Civic & landscape art	Landschaftsgestaltung, Raumplanung
ddc:720	Architecture	Architektur
ddc:730	Plastic arts Sculpture	Plastik, Numismatik, Keramik, Metallkunst
ddc:740	Drawing & decorative arts	Zeichnung, Kunsthandwerk

setSpec	setName	Deutschsprachige Beschreibung
ddc:741.5	Comics, Cartoons	Comics, Cartoons, Karikaturen
ddc:750	Painting & paintings	Malerei
ddc:760	Graphic arts Printmaking & prints	Grafische Verfahren, Drucke
ddc:770	Photography & photographs	Fotografie, Computerkunst
ddc:780	Music	Musik
ddc:790	Recreational & performing arts	Freizeitgestaltung, Darstellende Kunst
ddc:791	Public performances	Öffentliche Darbietungen, Film, Rundfunk
ddc:792	Stage presentations	Theater, Tanz
ddc:793	Indoor games & amusements	Spiel
ddc:796	Athletic & outdoor sports & games	Sport
ddc:800	Literature & rhetoric	Literatur, Rhetorik, Literaturwissenschaft
ddc:810	American literature in English	Englische Literatur Amerikas
ddc:820	English & Old English literatures	Englische Literatur
ddc:830	Literatures of Germanic languages	Deutsche Literatur
ddc:839	Other Germanic literatures	Literatur in anderen germanischen Sprachen
ddc:840	Literatures of Romance languages	Französische Literatur
ddc:850	Italian, Romanian, Rhaeto-Romanic literatures	Italienische, rumänische, rätoromanische Literatur
ddc:860	Spanish & Portuguese literatures	Spanische und portugiesische Literatur
ddc:870	Italic literatures Latin	Lateinische Literatur
ddc:880	Hellenic literatures Classical Greek	Griechische Literatur
ddc:890	Literatures of other languages	Literatur in anderen Sprachen
ddc:900	Geography & history	Geschichte
ddc:910	Geography & travel	Geografie, Reisen
ddc:914.3	Geography & travel Germany	Landeskunde Deutschlands
ddc:920	Biography, genealogy, insignia	Biografie, Genealogie, Heraldik
ddc:930	History of the ancient world	Alte Geschichte, Archäologie
ddc:940	General history of Europe	Geschichte Europas
ddc:943	General history of Europe Central Europe Germany	Geschichte Deutschlands
ddc:950	General history of Asia Far East	Geschichte Asiens
ddc:960	General history of Africa	Geschichte Afrikas
ddc:970	General history of North America	Geschichte Nordamerikas
ddc:980	General history of South America	Geschichte Südamerikas
ddc:990	General history of other areas	Geschichte der übrigen Welt

Das folgende Beispiel zeigt einen Ausschnitt aus einer möglichen Antwort eines Datenproviders auf eine **ListSets**-Anfrage, der den Empfehlungen über die inhaltliche Gliederung folgt.

```
<set>
    <setSpec>ddc</setSpec>
    <setName>DDC classified objects</setName>
</set>
<set>
    <setSpec>ddc:004</setSpec>
    <setName>Data processing Computer science</setName>
</set>
<set>
    <setSpec>ddc:610</setSpec>
    <setName>Medical sciences Medicine</setName>
</set>
```

2.3 Gliederung nach dem Publikationstyp

Als zweite Möglichkeit der Gliederung eines OAI-kompatiblen Archivs mittels Sets wird eine Einteilung der Objekte nach deren formalem Publikationstyp empfohlen.

Tabelle 2 zeigt die verwendeten Publikationstypen (Spalte 3) mit den entsprechenden Bezeichnern (Spalte 1) und Beschreibungen (Spalte 2) der dazugehörigen Sets.

Tabelle 2: Bezeichnung und Beschreibung der Sets nach dem Publikationstyp

SetSpec	SetName	Deutschsprachige Beschreibung
pub-type	Objects having a formal publication type	Objekte mit einem formalen Publikationstyp
pub-type:monograph	Books, Monographs	Bücher, Monographien
pub-type:article	Journal Articles	Zeitschriftenartikel
pub-type:dissertation	Dissertations and Professional Dissertations	Dissertationen und Habilitationen
pub-type:masterthesis	Diploma Theses	Diplomarbeiten
pub-type:report	Reports	Berichte
pub-type:paper	Papers	Papers
pub-type:conf-proceeding	Conference Proceedings	Tagungs- und Konferenzbeiträge
pub-type:lecture	Lectures	Vorlesungen
pub-type:music	Music	Musik
pub-type:program	Programs / Software	Programme / Software
Pub-type:play	Plays	Schauspiele / Theaterstücke
Pub-type:news	News	Nachrichten
Pub-type:standards	Standards	Standards

Der nachfolgende Ausschnitt aus einer XML-Datei stellt das Beispiel eines Teils einer Antwort auf die OAI-Anfrage **ListSets** dar.

```
<set>
    <setSpec>pub-type</setSpec>
    <setName>Documents having a formal publication-type</setName>
</set>
<set>
    <setSpec>pub-type:monograph</setSpec>
    <setName>Books, Monographs</setName>
</set>
<set>
    <setSpec>pub-type:dissertation</setSpec>
    <setName>Dissertations and Professional Dissertations</setName>
</set>
```

2.4 Gliederung nach dem Dokumenttyp

Als dritte Kategorie für die Strukturierung eines Datenproviders wird der Dokumenttyp der digitalen Objekte zur Unterscheidung verwendet. In Tabelle 3 sind die unterstützten Dokumenttypen (Spalte 3) mit den jeweiligen Bezeichnern (Spalte 1) und Beschreibungen (Spalte 2) der dazugehörigen Sets dargestellt.

Tabelle 3: Bezeichnung und Beschreibung der Sets nach dem Dokumenttyp

SetSpec	SetName	Deutschsprachige Beschreibung
doc-type	Objects having a formal document type	Objekte mit einem formalen Dokumenttyp
doc-type:text	Text	Text
doc-type:notes	Notes	Noten
doc-type:image	Images	Bilder
doc-type:audio	Audio files	Audiodateien
doc-type:video	Video files	Videodateien
doc-type:multimedia	Multimedia files	Multimediadateien
doc-type:data	Data	Daten
doc-type-binary	Binary data, (executable) programs	Binärdaten, (ausführbare) Programme

Die nachfolgende XML-Sequenz ist ein Ausschnitt aus einer möglichen OAI-Antwort auf die ListSets-Anfrage an einen Datenprovider, der Metadaten über Videodateien besitzt.

```
<set>
    <setSpec>doc-type</setSpec>
    <setName>formal document-type</setName>
</set>
<set>
    <setSpec>doc-type:video</setSpec>
    <setName>Video files</setName>
</set>
```

3 Weitere Empfehlungen

3.1 Dublin Core und andere Metadatenformate

Jeder Datenprovider soll entsprechend dem OAI-PMH seine Metadaten zumindest nach dem unqualifizierten Dublin-Core-Standard anbieten. Dies stellt keinerlei Beschränkungen bezüglich der Verfügbarmachung beliebiger anderer Metadatenformate dar. Um einen qualitativ hochwertigen Metadatenaustausch zwischen Daten- und Serviceprovidern zu erreichen, ist es sinnvoll, sich – zumindest innerhalb von Communities – auf die Verwendung spezialisierter Metadatenformate zu einigen. Ähnlich wie bei der Festlegung von Sets, wie sie in diesem Papier empfohlen werden, ist es auch auf diesem Gebiet notwendig, entsprechende Einigungen zu erreichen. Diese haben allerdings einen weniger übergreifenden Charakter als die vorliegenden Empfehlungen zur Set-Hierarchie. Stattdessen müssen Empfehlungen zu Metadatensätzen und spezialisierten Set-Definitionen auf der Ebene von (fachlichen) Communities erarbeitet und verbreitet werden.

Das Metadatenformat Dublin Core definiert 15 Elemente, die im Einzelnen aufgeführt werden. Diese Felder können praktisch mit beliebigem Inhalt gefüllt werden. Um eine bessere Interoperabilität zu gewährleisten, sollte man sich jedoch an den Empfehlungen der Dublin-Core-Arbeitsgruppe orientieren.

Diese Empfehlungen werden in Tabelle 4 in verkürzter Form wiedergegeben, die ausführlichen Beschreibungen finden sich in den Empfehlungen der Arbeitsgruppe[7].

[7] http://www.ietf.org/rfc/rfc2413.txt

Tabelle 4: Empfehlungen für den Inhalt der DC-Elemente

Dublin-Core-Element	Empfehlungen
Title	beliebiger Text
Author	beliebiger Text
Subject	beliebiger Text
Description	beliebiger Text
Publisher	beliebiger Text
Contributor	beliebiger Text
Date	Empfohlene Praxis ist ein Unterformat von ISO 8601 [W3CDTF][8] und schließt Datumsangaben der Form YYYY-MM-DD ein.
Type	kontrolliertes Vokabular (z.B. das *DCMI Type Vocabulary* [DCT1][9])
Format	kontrolliertes Vokabular (z.B. die Liste der *Internet Media Types* [MIME][10], in der digitale Medienformate definiert werden)
Identifier	Zu den häufig verwendeten formalen Identifikationssystemen zählen der *Uniform Resource Identifier* (URI), der *Digital Object Identifier* (DOI) und die *International Standard Book Number* (ISBN). Der Inhalt dieses Feldes ist jedoch nicht auf diese Systeme beschränkt.
Source	beliebiger Text
Language	RFC3066[11] in Verbindung mit ISO639[12], welches zwei oder drei Buchstaben für die Sprachen benutzt. Beispiele sind "en" oder "eng" für englisch.
Relation	beliebiger Text
Coverage	beliebiger Text
Rights	beliebiger Text. Empfehlenswert ist jedoch die *Creative Commons License*, die automatisch ausgewertet werden kann.

Mittels des *DC Checker*[13] lässt sich überprüfen, ob die Metadaten, die man über seinen eigenen Datenprovider anbietet, diesen Empfehlungen entsprechen.

[8] http://www.w3.org/TR/NOTE-datetime

[9] http://dublincore.org/documents/dcmi-type-vocabulary/

[10] http://www.isi.edu/in-notes/iana/assignments/media-types/media-types

[11] http://www.ietf.org/rfc/rfc3066.txt

[12] http://www.loc.gov/standards/iso639-2/langhome.html

[13] http://harvest.physik.uni-oldenburg.de/dc/dcchecker.php

Impressum

Diese Empfehlungen wurden von der DINI-Arbeitsgruppe zu OAI erarbeitet. Sie sind auf dem DINI Server unter http://www.dini.de/ veröffentlicht. Für kritische Hinweise, Korrekturvorschläge und ergänzende Bemerkungen sind wir dankbar. Zur besseren Koordination einer möglichen Diskussion bitten wir Sie, Ihre Anmerkungen per E-Mail an die DINI-Geschäftsstelle (*gs@dini.de*) zu senden.

Name	*Vorname*	*Institution*	*E-Mail*
Diekmann	Bernd	Carl-von-Ossietzky-Universität Oldenburg, Bibliotheks- und Informationssystem	diekmann@bis.uni-oldenburg.de
Dobratz	Susanne	Humboldt-Universität zu Berlin, Universitätsbibliothek	dobratz@cms.hu-berlin.de
Dr. Klotz-Berendes	Bruno	Hochschulbibliothek Münster	klotz-berendes@fh-muenster.de
Müller	Uwe	Humboldt-Universität zu Berlin, Computer- und Medienservice	u.mueller@cms.hu-berlin.de
Scholze	Frank	Universität Stuttgart, Universitätsbibliothek	scholze@ub.uni-stuttgart.de
Dr. Stamerjohanns	Heinrich	Institute for Science Networking, Carl-von-Ossietzky-Universität Oldenburg	stamer@uni-oldenburg.de