



IFLA  
2005  
OSLO

## World Library and Information Congress: 71th IFLA General Conference and Council

### "Libraries - A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

*Conference Programme:*

<http://www.ifla.org/IV/ifla71/Programme.htm>

08 June, 2005

**Code Number:**

**060-G**

**Meeting:**

**150 SI - ICABS (IFLA/CDNL Alliance for Bibliographic Standards)**

### **Metadaten-Standards für digitale Ressourcen Was wir haben und was wir brauchen**

**Sally H. McCallum**

Library of Congress  
Washington DC, USA

#### **Abstract:**

*Ein wesentlicher Bestandteil für die erfolgreiche Erhaltung digitaler Ressourcen werden Metadaten sein, die es ermöglichen, automatische Archivierungsprozesse ablaufen zu lassen. Die Menge der digitalen Einzeldokumente schließt manuelle Bearbeitung aus, und die Tatsache, dass es sich um elektronische Ressourcen handelt, prädestiniert sie für computergesteuerte Archivierungsaktivitäten. Während des letzten Jahrzehnts gab es eine Anzahl von Versuchen mit Dokumentenservern, die von unterschiedlichen Ansätzen ausgegangen sind, verschiedene Datenmodelle entwickelt und angewendet haben, und hauptsächlich unser Verständnis weitergebracht haben.*

*Dieser Vortrag berichtet über die neue Initiative PREMIS, die auf den bisherigen Konzepten und Erfahrungen aufbaut. Sie verdient sorgfältiges Testen, um zu sehen, ob die identifizierten Metadaten allgemein verwendet werden und eine Grundlage für detaillierte Metadaten werden können. Und wie viel mehr wird für Erhaltungsaktivitäten gebraucht? Initiativen für zusätzliche technische Metadaten und Dokumentformat-Registries werden ebenfalls angesprochen.*

## **Kern-Metadaten zur Erhaltung: PREMIS**

### *Ursprung*

Die Preservation Metadata Implementation Strategies (PREMIS)-Projektgruppe<sup>1</sup> ist aus den Erfahrungen des letzten Jahrzehnts erwachsen. Es gibt bedeutende Arbeiten über Archivserver-Systeme im Bibliothekswesen und insbesondere bei den Partner-Institutionen von ICABS und deren Mitarbeitern. Diese Arbeit hat zwangsläufig das Design einiger formeller und informeller Datenmodelle und die Identifizierung von Datenelementen für die Archivierungsfunktion beinhaltet, obwohl sie oftmals breiter angelegte Ziele hatte als die Erhaltung, indem sie den Fokus auf Zugänglichkeit und Verbreitungsfragen legte. Einige dieser Projekte waren das Projekt Networked European Deposit Library (NEDLIB) unter Federführung der Nationalbibliothek der Niederlande, das Project CURL Exemplars in Digital Libraries (CEDARS) aus Großbritannien, das Pandora-Projekt der National Library of Australia und verschiedene institutionelle Initiativen wie die von OCLC, die Erfahrung der Library of Congress mit der National Digital Library und andere.

Interessanterweise bezogen sich all diese Projekte an einem Punkt auf das Open Archival Information System (OAIS)-Referenzmodell<sup>2</sup>, das zum ersten Mal für Weltraum-Datensysteme artikuliert wurde und später zum ISO-Standard wurde (ISO 14721). Das OAIS-Modell hat einen vereinheitlichenden Einfluss auf die Untersuchungen des letzten Jahrzehnts gehabt, selbst wenn es auch nur eine Sprache auf hohem Niveau geliefert hat, um die Diskussionen zu unterstützen. Die Archivierungs-, Vorlage- und Verbreitungs-Informationspakete (AIP, SIP und DIP) als konzeptionelle Grundbestandteile für die Einrichtung digitaler Archive sind allgemein verständlich. Diese Informationspakete bestehen aus 4 Teilen, die sich auf das Informationsobjekt beziehen, das behandelt wird: die Inhaltsinformation, die Pack-Information, die beschreibende Information und – worauf wir uns beschränken – die Archivierungsinformation. Im Rahmen eines von OCLC und RLG finanzierten Projektes wurden im Jahr 2002 die Modelle und Metadaten, die in den o. g. Projekten spezifiziert wurden, in einen Zusammenhang gebracht und in das breite Konzept des OAIS-Referenz-Modells eingegliedert.<sup>3</sup> Die erste Aufgabe der PREMIS-Arbeitsgruppe war es daher, diese Ergebnisse wieder aufzugreifen und sie mittels eines Daten-Verzeichnisses in ein Set von realisierbaren Datenelementen zu übersetzen.

### *Ziele*

Das PREMIS-Projekt war das mehrjährige Bemühen einer Arbeitsgruppe unter Beteiligung von Institutionen mit nennenswerten Implementationen aus der ganzen Welt. Vertreter aus Australien, Neuseeland, den Vereinigten Staaten, Großbritannien, den Niederlanden und Deutschland trugen auf verschiedene Art bei, und manche mussten zu früher Stunde aufstehen, um an den wöchentlichen Telefonkonferenzen teilnehmen zu können. Die für ein Jahr geplante Arbeit hat zwei Jahre gedauert, aber das Ergebnis ist ein höchst präzises Set an Elementen, das als Grundlage für Implementationen dienen kann.

Die Arbeiten hatten mehrere miteinander verwandte Ziele, die alle praktischer Natur waren, und beabsichtigten, den Konzepten eine Anwendungsgrundlage zu geben. Die ursprünglichen Ziele beinhalteten die Identifizierung eines Kern-Sets von Metadaten und die Entwicklung eines Data-Dictionary für diese Metadaten, was jetzt beides erfolgreich fertig gestellt wurde. Experimentieren mit dem Data-Dictionary wird die beste Methode zur Formulierung alternativer Anwendungsstrategien sein, das dritte Ziel. Die Endziele, Pilotversuche mit dem Data-Dictionary und kooperative

Programme, die auf den Kernelementen basieren, sollen der gegenwärtigen Arbeit folgen.

### *Umfrage*

Die Arbeit begann mit der Befragung einer Reihe von Anwendungen digitaler Archivserver-Projekte zur Identifizierung gegenwärtiger Verfahren und Trends für digitale Projekte. Die Umfrage ergab 48 Antworten aus 13 Ländern, eine gute Quote für ein sich erst entwickelndes Arbeitsfeld. Die allgemeinen Schlussfolgerungen aus der Umfrage<sup>4</sup>, die als Information für die Arbeit am Data-Dictionary diente, welche parallel zur Umfrage und im Anschluss daran weiterlief, können folgendermaßen zusammengefasst werden:

- Das OAIS-Referenzmodell findet als Rahmen und Ausgangspunkt für das Design von Speichern breite Anwendung.
- Es ist eine weit verbreitete Praxis, Metadaten in den Speichersystemen redundant zu speichern: in einer XML- oder relationalen Datenbank für schnelles Retrieval und flexible Meldung sowie mit dem Inhaltsobjekt selbst zur Definition und zukünftigen Archivierung.
- Der Metadata Encoding and Transmission Standard (METS) wird umfassend angewendet zur Verschlüsselung des breiteren Spektrums an Metadaten, die für digitale Objekte benötigt werden, einschließlich der Archivierungsmetadaten, gemeinsam mit MIX (Metadata for Images in XML), das innerhalb von METS für technische Metadaten zu Bildern benutzt wird.
- Der gegenwärtige Trend ist es, das Original zu erhalten und verschiedene normalisierte und/oder migrierte Versionen des Inhaltsobjektes zu speichern, jeweils mit den dazugehörigen Metadaten.
- Die Verfolgung mehrerer Strategien sogar innerhalb einer Institution ist in einem solch experimentellen und sich entwickelnden Bereich nicht ungewöhnlich.

Zusätzlich zeigte die Umfrage, dass für Metadaten, die sich auf verschiedene Objekttypen beziehen (Bit-Streams, Dateien, Sammlungen, logische Objekte etc.) eine Reihe von Unterscheidungen gemacht wurden, und häufig wurden Informationen angegeben, die die Beziehungen zwischen Objekten anzeigen.

### *Data-Dictionary*

Anknüpfend an das vorherige Rahmen-Projekt (indirekt auch an die verschiedenen Hauptprojekte des letzten Jahrzehnts) und an die Information aus der Umfrage zu den digitalen Archiven wurde von der PREMIS-Arbeitsgruppe das Data-Dictionary der Kernelemente entwickelt.<sup>5</sup> In den frühen Projektphasen wurden einige Entscheidungen getroffen, die für die Anwendbarkeit wichtig sind.

Kerndatenelemente wurden von der Arbeitsgruppe gedeutet als „Dinge, die die meisten Langzeitarchivsysteme wahrscheinlich wissen müssen, um die digitale Archivierung zu unterstützen.“<sup>6</sup> Die Gruppe hat absichtlich einige sehr bekannte Aspekte der Erhaltung, wie z. B. technische Metadaten für verschiedene Medien, nicht behandelt. Nur technische Metadaten, die gewöhnlich auf alle Datenformate anwendbar sind, wurden durch die Arbeitsgruppe für PREMIS verfolgt.

Ein weiterer wichtiger Faktor, den die Arbeitsgruppe einbezogen hat, war, dass die spezifischen Metadaten automatisch geliefert und benutzt werden können müssen, soweit dies möglich ist. Dies führte dazu, dass man Werten aus autorisierten Listen gegenüber textlichen Beschreibungen den Vorzug gab. Dies passt auch zu der Absicht der Arbeitsgruppe, die Anwendung des Data-Dictionary unabhängig zu machen. Wie die Umfrage zeigte, sind bereits Archivsysteme in Produktion und bei denjenigen, die in der Planungsphase sind, kann die Systemumgebung, in der sie sich aufhalten, spezielle Eigenschaften haben. Die PREMIS-Kernelemente, die für das Archivsystem vorliegen, müssen nicht ausdrücklich in ihm gespeichert sein. Die Elemente könnten in Hilfssystemen gespeichert sein, könnten in den Betriebsanweisungen enthalten sein, die das Archivsystem verwendet oder könnten in einer lokalen Datenbank bzw. in einem lokalen Format gespeichert werden. Der wichtige Punkt dabei ist, dass im Falle eines Austausches die Kerndaten für die Konvertierung in einen Standard vorhanden sind. Oder dass die Daten vorhersehbar für jede Software verfügbar sind, die das Archivsystem auswählen könnte und die davon ausgeht, dass die PREMIS-Kernelemente zugänglich sind. Systeme müssen nicht neu implementiert oder speziell designed werden, um den PREMIS-Kern in einem Standard-Format bereitzuhalten. Dies brachte die Arbeitsgruppe dazu, in dem Data-Dictionary „semantische Einheiten“ anstelle von „Metadaten-Elementen“ zu definieren.

### *Datenmodell*

Während dieser Vortrag zu kurz für eine detaillierte Beschreibung des Datenmodells ist, ist es wichtig, einige neue Merkmale besonders hervorzuheben. (Das Modell ist vollständig und gut erklärt im PREMIS report, s. Fußnote 5.)

Das Modell ist einfach. Es gibt nur fünf Arten von Entitäten: Objekte, Ereignisse, Agents, Rechte und die intellektuelle Entität selbst. Die Relevanz der Information, die in das Data-Dictionary eingeflossen ist, wurde sorgfältig unter die Lupe genommen. So werden z. B. beschreibende Metadaten, die den intellektuellen Inhalt beschreiben, der ein Buch, eine Karte, Web-Site usw. sein kann, den vielen, bereits existierenden Standards wie MARC, MODS (Metadata Object Description Standard) und DC (Dublin Core) überlassen. Ebenso werden detaillierte Daten über Agents MARC, MADS (Metadata Authority Description Standard), vCard und anderen Standards überlassen. Rechte-Daten werden auf diejenigen beschränkt, die sich auf die Erlaubnis zu Erhaltungsaktivitäten beziehen, weil Zugriffs- oder Verbreitungsrechte für das Objekt nicht zum Kern der Archivierungsaktivitäten gehören. Detaillierte technische Metadaten und Medien- sowie Hardware-Dokumentationen sind nicht eingeschlossen, sondern den Formatexperten zur Bestimmung überlassen.

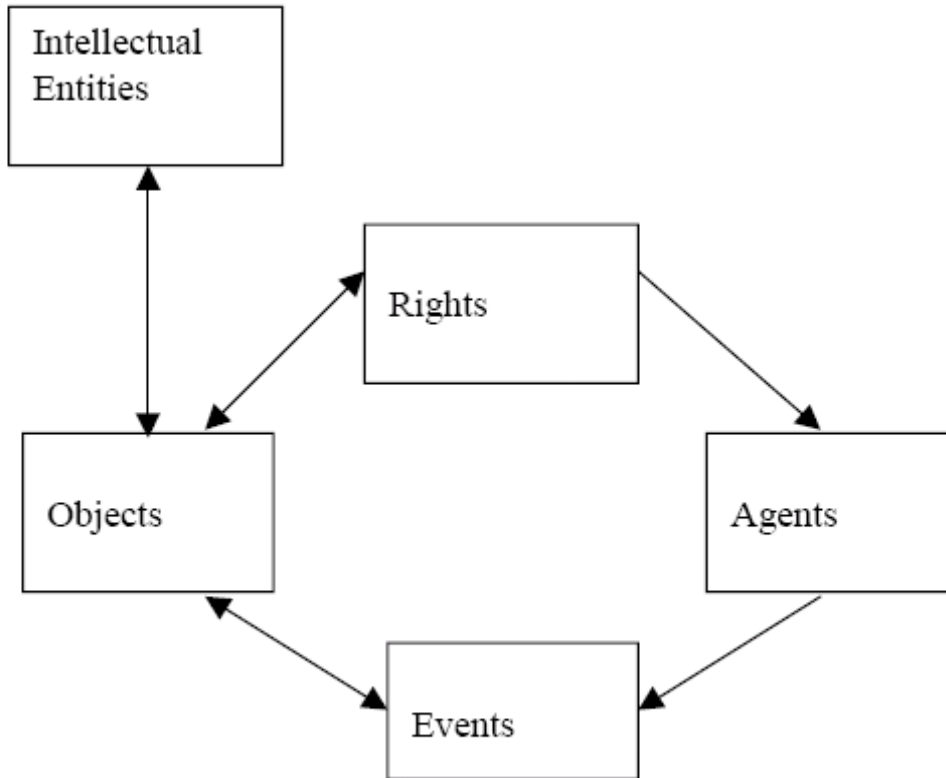


Figure 1: Basic PREMIS data model

Semantische Einheiten für Objekte, das Kernkonzept im Modell, können auf drei Ebenen festgelegt werden, welche Flexibilität bieten, Informationen einzuschließen auf den Ebenen, die zum Material passen, und für den Betrieb eines Langzeitspeichers. Diese Ebenen sind *Bit-Stream*, der ein Bestandteil der nächsten Ebene ist, der *Datei* (oder dem *Daten-Strom*). Ein Set von Dateien, die für eine vollständige Wiedergabe einer intellektuellen Entität benötigt werden, bildet die höchste Ebene, die *Darstellung*.

Die Entität Ereignis, die auf das Objekt bezogene Aktionen dokumentiert, ist ein wichtiger Bestandteil des Modells. Eine große Vielfalt an Aktionen einschließlich Veränderungen des Objekts, durchgeführte Validitäts- und Integritäts-Checks, sogar Anfragen zur Verbreitung oder zu Berichten beeinflussen die Erhaltung von digitalem Material.

Ereignisse hängen auch häufig mit Beziehungen zusammen, weil ein Ableitungs-Ereignis ein weiteres Objekt produziert, und es normalerweise wichtig ist, die Beziehung zwischen den Objekten für Archivierungszwecke aufzuzeichnen. Das Data-Dictionary liefert mehrere semantische Einheiten für Beziehungen, im Zusammenhang mit der Aufzeichnung von Informationen über abgeleitete und strukturelle Beziehungen, Abhängigkeiten und andere Beziehungen.

Ein wichtiger Aspekt des Datenmodells ist das, was die Arbeitsgruppe das 1:1-Prinzip genannt hat. Neue Objekte, die aus bestehenden Objekten geschaffen wurden (Kopien, Versionen, Transformationen usw.) werden als neue Objekte

behandelt und mit dem „alten“ Objekt durch die Ereignis- und Beziehungs-Information verlinkt. Eines der Ergebnisse der Befragung war, dass Speicher oft mehrfache Kopien eines Objekts aufbewahren, und dass es für Archivierungszecke wichtig ist, dass die Daten zu jedem Objekt vollständig sind. Deshalb stellt die Beziehungsinformation das Bindeglied dar, ohne die vollständige Aufzeichnung von Archivierungsinformationen über die Ableitung zu schmälern oder kompliziert zu machen. Während ein Archivsystem intern Datenbäume aufbauen kann, um Datenredundanz zu verringern, muss das Archivsystem für den Austausch in der Lage sein, ein unabhängiges Objekt mit allen Archivierungs-Metadaten weiterzuleiten.

### *Nächster Schritt: Testen*

PREMIS ist eine sorgfältig ausgelotete, internationale Zusammenarbeit, die ein Data-Dictionary für Metadaten geschaffen hat, mit dem Potenzial, den Standardaustausch von Archivierungsinformation mit digitalem Material aus elektronischen Archiven zu ermöglichen. Es erzwingt keine spezifische Architektur von den Speichern, liefert aber eine Anleitung für ein Kernset von Archivierungsmetadaten. Trotz der weltweiten Beteiligung, wurde das PREMIS-Projekt von OCLC und RLG finanziert, und die Library of Congress hat die Verantwortung für die offizielle Website während der nächsten Phase übernommen.<sup>7</sup> Alle Projektdokumente und Nachrichten sind auf dieser Site erhältlich.

Die Endziele des Projekts, eine Testumgebung für das Data-Dictionary und Kooperationen zum Thema Metadaten, können nun geplant werden. Kürzlich wurde ein XML-Schema geschrieben für die semantischen Einheiten, die im Data-Dictionary identifiziert wurden.<sup>8</sup> Es muss angewendet und in neuen Projekten sowie für den Austausch getestet werden. Dennoch hofft man, dass bereits existierende Anwendungen von Archivsystemen oder Projekte, die mit einer speziellen Architektur geplant sind, ebenfalls an der Testumgebung teilnehmen werden, indem sie ihre Metadaten implizit und explizit analysieren und den semantischen Einheiten des Data-Dictionary gegenüberstellen. Zwischenzeitlich werden das Data-Dictionary und das XML-Schema stabil gehalten, werden aber Gegenstand laufender Revisionen sein, sobald Erfahrungen aus der Testumgebung gewonnen wurden.

### **Weitere Teile des Puzzles**

Wie oben erwähnt, gibt es andere Teile der notwendigen Archivierungsmetadaten für digitale Medien, die nicht von der PREMIS-Arbeitsgruppe festgelegt wurden – zum Beispiel komplexe Rechte-Metadaten und detaillierte technische Metadaten, einschließlich der Informationen zum digitalen Format.

### *Rechte-Metadaten*

Rechte-Metadaten wurden für PREMIS eng definiert, und man könnte einwenden, dass gewisse Zugangs- und Verbreitungsinformationen für Archivierungszwecke bedeutend sind. Eine Reihe von Initiativen jedoch sieht die Sprache für die Rechte und Fragen der Datentransferstandards als zum Zugriff und der Verbreitung gehörend an. Die Arbeit von Indecs der Europäischen Union, die Bemühungen der Verlegerverbände um ONIX und die Electronic Rights Management Initiative (ERMI) der Digital Library Federation (DLF) stellen einige bedeutende Untersuchungen dar.

## *Technische Metadaten*

Die PREMIS-Umfrage förderte zutage, dass viele Archivsysteme METS anwenden, um ihre digitalen Objekt-Metadaten zu bündeln, und dass es eine Vielzahl von Art und Menge der gehaltenen technischen Metadaten gab, abhängig davon, was das Archivsystem automatisch einsammeln konnte. Der eine Bereich, in dem Standardisierungsarbeit einen großen Fortschritt gemacht hatte, ist der Bereich mit Metadaten für Bildquellen. Ein Standard-Data-Dictionary wurde 2002 vom NISO für Testzwecke fertig gestellt<sup>9</sup>. MIX, das Schema, das METS ergänzt, und das auf dem NISO-Data-Dictionary basiert, ist dennoch schon weit verbreitet.<sup>10</sup> Das schnelle Aufgreifen dieses Standards und Schemas zeigt, dass Archivsysteme sehr an Standards und Anleitungen für detaillierte technische Informationen interessiert sind. Für detaillierte technische Metadaten muss die bibliothekarische Welt mit jedem aufkommenden Industriestandard zusammenarbeiten oder ihn zumindest aufmerksam zur Kenntnis nehmen, weil diese Ebene der Metadaten von den Objekten ableitbar sein muss – noch mehr als die Information auf der PREMIS-Ebene. Die METS-Website weist auf verschiedene lokal entwickelte technische Metadaten-Schemas für verschiedene Materialtypen hin, die möglicherweise als Ausgangspunkt für breiter angelegte Bemühungen zur Entwicklung von Standards dienen können, die mit denen für Bilddaten vergleichbar sind.<sup>11</sup>

## *Format-Registries*

Ein zweiter, möglicherweise ebenso wertvoller Teil auf der Archivierungsmetadaten-Schiene ist der leichte Zugang zu den Besonderheiten des elektronischen Datenformats. Diese Informationen kann man manchmal auf den Websites der Firmen finden, die für verschiedenen Datenformate zuständig sind, falls solch eine Site existiert, aber das ist kein effizienter Weg, die Informationen zu erhalten. Aus der Archivierungsperspektive hilft das Wissen über die Datenformate bei der Validierung digitaler Objekte bei der Aufnahme oder für Integritäts-Checks, und es hilft, das Risiko zu evaluieren, das mit verschiedenen digitalen Formaten einhergeht, und zeigt geeignete Migrationswege für digitale Objekte auf. Das Verstehen des Datenformats kann auch hilfreich sein, Metadaten zu bestimmen, die aus dem digitalen Objekt abgeleitet werden können und so dabei helfen, PREMIS und Datenbanken mit detaillierten technischen Metadaten anzureichern.

Es gibt zwei bekannte Projekte zur Entwicklung von kontinuierlich aktualisierten gemeinschaftlich angelegten Verzeichnissen, aber es steht noch nicht fest, ob sie weitergeführt werden können. Eines der Projekte ist PRONOM der National Archives in Großbritannien.<sup>12</sup> Diese Registry hat als ein lokal zusammengestelltes Tool angefangen, das von den National Archives benötigt wurde, um beim Kampf gegen das Veralten von Software zu helfen, indem es die Migration von Dokumenten lenkt. Es wurde 2004 im Web zugänglich gemacht, und 2005 gab es eine stark verbesserte Ausgabe. Mit einem Schwerpunkt auf Staatsarchiven, ist diese Registry besonders leistungsfähig bei textorientierten Software-Formaten.

Ein zweites Projekt, das zu einem Proof-of-concept-Stadium gelangt ist, ist Global Digital Format Registry (GDFR), das aus einer von der DLF geförderten Konferenz im Jahr 2003 hervorgegangen ist.<sup>13</sup> Nachdem das Modell dieser Registry von Harvard-Mitarbeitern veröffentlicht wurde, wurde ein prototypischer Formatdienst an der University of Pennsylvania mit dem Namen Format Registry Demonstration (FRED) entwickelt.<sup>14</sup> Mit FRED können Entwickler von Archivsystemen

experimentieren, wie ein solcher Dienst nützlich sein könnte, welche Dienste er anbieten können sollte, wie er gepflegt werden könnte usw.

Dies ist ein Gebiet, das nicht besonders glamourös ist, aber es ist wichtig für die Archivierung aller Medien – und eine gemeinschaftliche Registry wäre für alle wirtschaftlich.

## Schlussfolgerung

Aufbauend auf frühere konzeptionelle Modelle und Anwendungserfahrungen entwickeln sich Schritt für Schritt Richtlinien und Standards für Metadaten zur Unterstützung von Langzeitarchivierungsaktivitäten. Wer ein Archiv aufbaut, muss nicht länger das Rad neu erfinden. Testen der PREMIS-Kernlemente, Beachtung der detaillierten technischen Anforderungen und die Zusammenarbeit bei einer Datenformat-Registry stehen heute auf der Tagesordnung für zukünftige Entwicklungen.

**Übersetzung:** Susanne Oehlschläger und Dr. Thomas Wollschläger, Die Deutsche Bibliothek

---

<sup>1</sup> Offizielle Website von PREMIS: [www.loc.gov/standards/premis](http://www.loc.gov/standards/premis)

<sup>2</sup> *Reference Model for an Open Archival Information System (OAIS)*. Washington, DC: Consultative Committee for Space Data Systems, 2002.

([ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf](http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf))

<sup>3</sup> *A Metadata Framework to Support the Preservation of Digital Objects*. Dublin, Ohio: OCLC Online Computer Library Center, 2002.

([www.oclc.org/research/projects/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/projects/pmwg/pm_framework.pdf))

<sup>4</sup> *Implementing Preservation Repositories for Digital Materials: Current Practice and Emerging Trends in the Cultural Heritage Community*. Dublin, Ohio: OCLC Online Computer Library Center, 2004.

([www.oclc.org/research/projects/pmwg/surveyreport.pdf](http://www.oclc.org/research/projects/pmwg/surveyreport.pdf))

<sup>5</sup> *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, May 2005*.

([www.oclc.org/research/projects/pmwg/premis-final.pdf](http://www.oclc.org/research/projects/pmwg/premis-final.pdf))

<sup>6</sup> Ibid., S. ix.

<sup>7</sup> Offizielle Web-Site von PREMIS: [www.loc.gov/premis/](http://www.loc.gov/premis/).

<sup>8</sup> Das PREMIS-Schema kann über folgende Adresse bezogen werden:

[www.loc.gov/standards/premis/schemas.html](http://www.loc.gov/standards/premis/schemas.html)

<sup>9</sup> *Data Dictionary - Technical Metadata for Digital Still Images, NISO Z39.87-2002/AIIM 20-2002*.

([www.niso.org/standards/resources/z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/z39_87_trial_use.pdf))

<sup>10</sup> MIX ist erhältlich unter: [www.loc.gov/mix](http://www.loc.gov/mix)

<sup>11</sup> siehe [www.loc.gov/mets](http://www.loc.gov/mets)

<sup>12</sup> [www.nationalarchives.gov.uk/pronom/](http://www.nationalarchives.gov.uk/pronom/)

<sup>13</sup> [hul.harvard.edu/gdfr/](http://hul.harvard.edu/gdfr/)

<sup>14</sup> [tom.library.upenn.edu/fred/](http://tom.library.upenn.edu/fred/)