

Der Metadata Encoding and Transmission Standard (METS) zur Beschreibung digitalisierter und genuin digitaler Objekte

LUPPRIAN, Karl-Ernst, The Metadata Encoding & Transmission Standard (METS) for the description of digitized and born digital objects. Atlanti, Vol. 17, N. 1-2, Trieste 2007, pp. 97-100.

Original in German, abstract in English, Italian and Slovenian, summary in English

METS is a format defined with XML Schema for the encoding of digital collections with metadata. METS contains elements for structuring digital objects and their linking to descriptive and administrative metadata. This standard can be used for digitized and born digital archival records and supports the presentation in an electronic environment.

The paper gives an overview about METS and demonstrates its use in the project DAOFIND of the German Federal Archives.

LUPPRIAN, Karl-Ernst, Gli standard di codifica e trasmissione METS per la descrizione di oggetti digitalizzati o in formato nativo digitale. Atlanti, Vol. 17, N. 1-2, Trieste 2007, pp. 97-100.

METS è un formato definito con schema XML per la codifica di collezioni digitali con metadati. METS contiene elementi per strutturare oggetti digitali ed il loro linking a metadati descrittivi ed

In den letzten Jahren hat sich die Präsenz von Archiven im Internet erheblich verstärkt. Wie bei den Bibliotheken, die schon früher begonnen hatten, ihre Kataloge online zu stellen, wurden - und werden - vor allem Findmittel in elektronischer Form weltweit nutzbar gemacht. Hierzu haben Standards und standardisierte Werkzeuge wesentlich beigetragen. So wurde der Verzeichnungsstandard ISAD(G)¹ in EAD (Encoded Archival Description)² abgebildet, der Standard für die Beschreibung von Registraturbildnern ISAAR(CPF) in EAC (Encoded Archival Context)³, und für die Beschreibung archiverischer Einrichtungen steht EAG (Encoded Archival Guide)⁴ zur Verfügung. Da alle diese Kodierungen mit XML realisiert wurden, können - um ein Beispiel zu nennen - auch Findbücher, die nicht in EAD erstellt wurden, aber zu ISAD(G) kompatibel sind und in XML vorliegen, über ein Mapping nach EAD konvertiert werden. Dies ist z.B. bei den so genannten MidosaXML-Findbüchern der Fall, denen wir bei der Betrachtung des Projekts <daofind> begegnen werden.

Mit der Bereitstellung archiverischer Findmittel im Web sind die Wünsche der historischen Forschung jedoch nicht befriedigt. Schon vor Jahren forderte man „das Archiv auf dem Schreibtisch des Historikers“, also den Zugriff nicht nur auf die Findmittel, sondern auch auf die Bestände selbst.

Die Digitalisierung von Archivgut ist vor allem ein organisatorisches und mehr noch ein finanzielles Problem und nicht so sehr eine Frage der Technik. Dennoch ist es nicht damit getan, ein Archivale zu digitalisieren, die Images ins Web zu stellen und mit der Verzeichnungseinheit im Findmittel zu verlinken. Zwei Beispiele sollen dies verdeutlichen.

1. Eine mittelalterliche Urkunde: Ein Forscher ist schon zufrieden, wenn die Auflösung des Images das Lesen des Textes erlaubt. Ein anderer hingegen möchte paläographische Untersuchungen durchführen; dafür benötigt er ein Image in hoher Auflösung, besser noch die Möglichkeit einer abgestuften Vergrößerung (Zoom). Die Webseite muss also entsprechende Images und Werkzeuge bereit stellen.

2. Ein moderner Verwaltungsakt: Der Forscher will durch den Text scrollen, schwer zu lesende Passagen vergrößern und jederzeit wissen, an welcher Stelle des Aktes er sich befindet. Zusätzlich zu den im vorigen Beispiel genannten Eigenschaften muss die Webseite

1. ISAD(G) und ISAAR(CPF) findet man unter www.ica.org -> Download Center.

2. Siehe www.loc.gov/eac/

3. Siehe www.library.yale.edu/eac/

4. Siehe Blanca Desantes, *The Encoded Archival Guide (EAG) DTD and the Censo-Guía de los Archivos de España e Iberoamérica Project. An Electronic Guide to Spanish and Iberian American Archives*. In: «Journal of Archival Organization» 3, Heft 2/3, S. 23-38.

Bildunterschriften mit den jeweiligen Blatt- oder Seitenangaben liefern.

Des Weiteren müssen alle Elemente (Verzeichnungsdaten, Images) und die für die Darstellung benötigten Eigenschaften in einer von der Webpräsentation unabhängigen, jederzeit bearbeitbaren Form gespeichert werden können. Hier bietet sich nun die Verwendung des Standards METS an.

METS - Der Standard

METS (Metadata Encoding and Transmission Standard)⁵ wurde auf Initiative der Digital Library Federation entwickelt, wobei auf Vorarbeiten des Projekts „The Making of America II“ (MOA2) aufgebaut werden konnte. Der Standard wird heute von der Library of Congress betreut.⁶ Er vereinigt Erschließungsangaben, administrative Daten und Metadaten zur Struktur digitaler Objekte in XML-Dokumenten und kann auch für den Austausch derartiger Objekte zwischen sammelnden Institutionen oder für deren Übermittlung an Benutzer verwendet werden. Im Kontext des Open Archival Information System (OAIS) Referenzmodells kann ein METS-Dokument als Ablieferungseinheit (Submission Information Package, SIP), Archivierungseinheit (Archival Information Package, AIP) oder Bereitstellungseinheit (Dissemination Information Package, DIP) verwendet werden.

Ein METS-Dokument besteht aus sieben Hauptabschnitten:⁷

1. Der **Kopfteil (METS Header)** enthält Metadaten, die das jeweilige METS-Dokument selbst beschreiben und nennt auch Bearbeiter oder Herausgeber des Dokuments. Ein Beispiel:

```
<metsHdr          CREATEDATE="2007-07-29T09:00:00
RECORDSTATUS="Complete">
  <agent           ROLE="CREATOR"
    TYPE="ORGANIZATION"
    <name>National Archives</name>
  </agent>
  <agent           ROLE="ARCHIVIST"
    TYPE="INDIVIDUAL"
    <name>John Doe</name>
  </agent>
</metsHdr>
```

2. **Erschließungsangaben (Descriptive Metadata)** können sowohl Verweise auf ein externes Dokument (z.B. ein EAD-Findbuch) wie auch in das METS-Dokument eingebettete Angaben oder beides enthalten.

3. **Verwaltungsangaben (Administrative Metadata)**. Dieser Abschnitt enthält Informationen zur Herstellung und Speicherung von Dateien, zu Urheberrechten und zur digitalisierten Vorlage. Ferner werden hier Angaben zur Herkunft der Digitalisate abgelegt (zu Masterdateien, Derivaten sowie zu Migrationen).

amministrativi. Questo standard può venir usato per documenti digitalizzati o in formato nativo digitale, e supporta la presentazione in un contesto elettronico.

L'articolo fornisce uno sguardo generale su METS e ne esplica l'uso nel progetto DAOFIND degli Archivi Federali Tedeschi.

LUPPRIAN, Karl-Ernst, Mega podatkovno kodiranje in standard prenosa (METS) za opis digitaliziranih in novo nastalih objektov. Atlanti, Zv. 17, Št. 1-2, Trst 2007, str. 97-100.

Mega podatkovno kodiranje in standard prenosa (METS) v XML sbemi je format za kodiranje digitalnih zbirke z mega podatki. METS kodiranje vsebuje tudi elemente za strukturiranje digitalnih objektov in njihovo povezavo z deskriptivnimi in administrativnimi mega podatki. Ta standard se bo uporabljaj za digitalizirano in novo digitalno arhivsko dokumentacijo in za podporo pri predstavitvi nekega elektronskega okolja. Razprava prinaša nekakšne pregled o MRTS in prikazuje njegovo uporabo v projektu DAOFIND v nemških zveznih arhivih.

SUMMARY

The online presentation of archival finding aids is strongly supported by international standards like EAD, EAC or EAG. But researchers want - beyond consulting the finding aids - to study archival content online. For organizing, administrating and linking of images to corresponding entries in finding aids the standard METS has proven very useful. In the project <daofind> of the German Federal Archives a set of tools (called "MEX Editors") was developed. These allow archivists to build their finding aids, prepare them for online presentation and adding digitized content without specific knowledge of XML or other technological specialties of the applied standards.

5. Kurzerläuterung in: A Glossary of Archival and Records Terminology, s.v. URL: www.archivists.org/glossary/

6. Offizielle Webseite: www.loc.gov/standards/mets/

7. Verkürzte Wiedergabe des von Angelika Menne-Haritz ins Deutsche übersetzten Texts „METS Overview and Tutorial“, www.bundesarchiv.de/imperia/md/content/daofind/metsuebersicht.pdf

4. Im **Dateienabschnitt (File Section)** werden alle Dateien, aus denen das digitale Objekt besteht, aufgelistet. Einzelne zusammengehörige Dateien können mit dem Element `<fileGrp>` (Dateigruppe) zusammengefasst werden, um z.B. verschiedene Versionen auseinander halten zu können.

5. Die **Strukturbeschreibung (Structural Map)** ist der zentrale Bestandteil jedes METS-Dokuments. Sie bildet den inneren Aufbau des digitalen Objekts ab und verknüpft die Elemente der Struktur mit den Dateien, aus denen der Inhalt des digitalen Objekts besteht, sowie mit deren Metadaten.

6. Der Abschnitt **Strukturverknüpfungen (Structural Links)** ermöglicht es den Erstellern von METS-Dokumenten, vorhandene Hyperlinks zwischen einzelnen Knoten des im Strukturabschnitts dargestellten hierarchischen Aufbaus des digitalen Objekts zu beschreiben - eine besonders für die Archivierung von Webseiten wichtige Funktion.

7. Der Abschnitt **Verhalten (Behavior)** des digitalen Objekts kann verwendet werden, um ausführbare Anweisungen für das Verhalten mit den Inhalten in METS-Objekten zu verknüpfen. So kann z.B. für die Anzeige einer bestimmten Bilddatei ein Viewer mit Parametern für die Steuerung der Darstellung aufgerufen werden.

METS im Projekt <daofind>

Das Projekt <daofind> (digital archives in online finding aids)⁸ ist beim deutschen Bundesarchiv angesiedelt und wurde von der Andrew W. Mellon Foundation mitfinanziert. Es lief vom November 2004 bis Juni 2006. Ziel des Projekts war es, eine Pilotanwendung für die kombinierte Nutzung der oben genannten Standards EAD, EAC und METS für einen verbesserten Online-Zugang zu Archivgut und den Prototypen für ein Editionswerkzeug für die Online-Präsentation auf der Basis von Open Source Software zu entwickeln.

EAD dient der Herstellung und Präsentation von strukturierten Online-Findmitteln. Man kann in ihnen anhand der Gliederung wie in einem Inhaltsverzeichnis navigieren, nach Begriffen suchen, Indizes nutzen und im Text blättern.

EAC verbessert die Nutzung von in den Findmitteln vorkommenden Namen von Personen, Orten und Organisationen, indem ihre Verwendung im Findmittel näher erläutert wird. EAC beschreibt also die Kontextinformationen in archivischen Erschließungsangaben auf der Ebene der Beständeübersicht oder der Findbücher.

METS erfasst und gliedert alle notwendigen Angaben zu digitalisiertem Archivgut und verknüpft die digitalen Objekte mit den zugehörigen Online-Findmitteln.

Auf der Basis der Entwicklungsumgebung Eclipse wurde eine Gruppe von Werkzeugen (MEX-Editoren) realisiert. Diese Editoren verwenden eine an archivische Arbeitsgewohnheiten angepasste

8. Die umfangreiche Webseite des Projekts findet sich unter www.daofind.de. Die folgende Darstellung ist eine Zusammenfassung.

Oberfläche. Sie nutzen Subsets der genannten drei Standards und bieten diese in der archivfachlichen Terminologie an. Das heißt, dass der Bearbeiter sich nicht mit den darunter liegenden XML-Tags oder gar mit dieser Technologie befassen muss. Die MEX-Editoren liegen zur Zeit in einer Zwischenversion (in englisch und deutsch) vor, die von der <daofind>-Webseite heruntergeladen werden kann. Lizenzgebühren werden nicht erhoben.

Anhand von drei Archivbeständen des Bundesarchivs kann man die Online-Präsentation, die auch 30.000 Images enthält, von der <daofind>-Webseite aus testen. Man recherchiert zunächst im Findmittel. Dort ist angegeben, ob zu einem Verzeichnungseintrag digitalisiertes Archivgut vorhanden ist. Anschließend gelangt man in eine Übersicht der angebotenen Images, die nicht Thumbnails enthält (diese wären unlesbar), sondern jeweils das erste Drittel des jeweiligen Images bzw. der Imagesequenz. Darin kann dann gescrollt werden.

Im Oktober 2006 hat das Bundesarchiv bei der Andrew W. Mellon Foundation die Förderung eines Fortsetzungprojekts <daofind+> beantragt, um u.a. die MEX-Werkzeuge weiter zu entwickeln und auch für Beständeübersichten, Fallakten und Nachlässe anwenden zu können.

Mit dem Projekt <daofind> wurde eine Lösung für die Erstellung und die multimediale Präsentation archivischer Findmittel und von Archivgut in digitaler Form gefunden, die auf internationalen Standards beruht, für die Kodierungen durchgehend den XML-Standard verwendet und lizenzfrei ist. Die Struktur der Findmittel richtet sich nach deutschen archivischen Traditionen⁹; eine Anpassung an andere nationale oder regionale Gewohnheiten dürfte nicht übermäßig schwierig sein.

9. Die Anwendung von EAD sichert jedoch die Recherchierbarkeit auf internationaler Ebene.