



UvA-DARE (Digital Academic Repository)

METS: Metadata Encoding & Transmission Standard

Brandsma, R.

Published in:
Informatie Professional

[Link to publication](#)

Citation for published version (APA):

Brandsma, R. (2007). METS: Metadata Encoding & Transmission Standard. *Informatie Professional*, 11(11), 35-37.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<http://dare.uva.nl>)

METS: Metadata Encoding & Transmission Standard

Als je een boek beschrijft in een bibliotheekcatalogus, kun je dat boek vervolgens dankzij een signatuur uit de kast pakken. Maar zodra je dat boek hebt gedigitaliseerd, beschik je niet meer over één object, maar over een reeks digitale bestanden. Hoe weet je nu welke bestanden samen dat boek vormen? Bij dat soort problemen kan METS helpen.

Renze Brandsma

Om digitale objecten te kunnen beheren en gebruiken, zijn andere en uitgebreidere metadata vereist dan voor een collectie van gedrukte werken. Bij een boek kan een bibliotheek volstaan met beschrijvende metadata en plaatsingsgegevens over dat boek als geheel, zonder het risico te lopen dat de afzonderlijke pagina's van dat boek onvindbaar worden.

Bij de digitale versie van dat boek is dat minder eenvoudig. Zonder *structurele metadata* die de scans van de afzonderlijke pagina's of de tekstbestanden aan elkaar koppelen, is het digitale werk van weinig waarde. En dan is er nog het bijkomende probleem dat de bestanden in de toekomst onleesbaar worden, omdat bijvoorbeeld het bestandsformaat niet langer gangbaar is. Daarom dient de bibliotheek voor intern beheer ook te beschikken over *technische metadata*. Die zijn nodig om periodiek de data te verversen en migreren en zo de duurzaamheid van de digitale werken te garanderen.

Complexe objecten

Een digitaal object kan bestaan

uit teksten, audio, video, images en data die onderling gerelateerd zijn en in een bepaalde volgorde of samenhang bekeken dienen te worden. Er is dan sprake van complexe of samengestelde digitale objecten (soms ook wel Scientific Publication Packages genoemd). Er zijn verschillende technologieën en standaarden om het beheer, de opslag en beschikbaarstelling van digitale complexe objecten te ondersteunen, waaronder RDF, METS en MPEG DIDL. In deze aflevering van De Standaard gaan we in op METS.¹

METS ofwel Metadata Encoding & Transmission Standard is een XML-documentformaat voor het coderen van metadata voor beheer, opslag en beschikbaarstelling van digitale bibliotheekobjecten en voor de uitwisseling van digitale objecten. Het is (nu nog) beperkt tot objecten die bestaan uit tekst, images, audio- en/of videobestanden. METS, een initiatief van de Digital Library Federation, bouwt voort op de resultaten uit het Making of America II-project (MOA2). Making of America werd in 1995 opgezet voor het di-

gitaliseren en ontsluiten van omvangrijke collecties primaire bronnen over de sociale geschiedenis van de Verenigde Staten.

Het METS-formaat bevat beschrijvende metadata over inhoud en context, administratieve metadata, bijvoorbeeld met betrekking tot rechten en toegangscontrole, en technische gegevens, bijvoorbeeld over de resolutie en structurele metadata. Eventueel kan ook de werkelijke content of het bestand zelf in METS opgenomen worden.

Technische componenten

METS bestaat uit een aantal technische componenten:

- 1 Het XML-schema.** Het basis-schema dat de regels geeft voor het opbouwen van een METS-bestand in XML. Een METS-bestand is opgebouwd uit zeven verschillende onderdelen (zie kader op pagina 37).
- 2 METS-profielen.** Dit zijn gedetailleerde toepassingen voor bijvoorbeeld specifieke documenttypen, of een toepassing binnen een bepaalde instelling. METS-profielen zijn bedoeld om

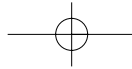
een klasse van METS-documenten in voldoende detail te beschrijven, zodat de vervaardigers en programmeurs richtlijnen hebben om documenten in die klasse te maken en te verwerken. In de richtlijnen kunnen de verplichte en aanbevolen elementen en inhoud worden vastgelegd. Een klasse kan een bepaald documenttype zijn, bijvoorbeeld audio. Instituten die een profiel hebben gemaakt kunnen deze laten registreren via de website van de Library of Congress. Dat kan pas nadat ze goedgekeurd zijn door de METS Editorial Board. Vervolgens zijn de profielen publiekelijk beschikbaar. Een voorbeeld van een profiel is de Library of Congress METS Profile for Audio Compact Discs.

- 3 METS externe schema's en gecontroleerde vocabulaires.** In METS is het mogelijk om te verwijzen naar allerlei willekeurige externe beschrijvende en administratieve metadata-schema's. Voor het gemak is een aantal bekende metadata-schema's al in een attribuutgroep in het METS-schema opgenomen. Voor de beschrijvende metadata zijn dat bijvoorbeeld Dublin Core, MODS, MARC, EAD, TEI en LOM. Voor administratieve metadata is dat bijvoorbeeld PREMIS (preservation metadata, ontwikkeld door OCLC-RLG PREservation Metadata: Implementation Strategies Working Group).²

Het onderhoud

De METS Editorial Board is verantwoordelijk voor het redactionele beheer, voor het primaire XML-schema, de profielen en de officiële documentatie. Ook zorgt de Editorial Board voor de promotie van het gebruik van de standaard en het onderhoud van de profielenregistratie. Voorts beveelt zij 'best practices' aan in het gebruik van de standaard.

Lees verder op pagina 37



DE STANDAARD

Voorbeelden van onderdelen uit een METS-document

1. METS header <metsHdr>

Een voorbeeld waar Jerome McDonough de vervaardiger is van het METS-document:

```
<metsHdr CREATEDATE="2003-07-04T15:00:00" RECORDSTATUS="Complete">
  <agent ROLE="CREATOR" TYPE="INDIVIDUAL">
    <name>Jerome McDonough</name>
  </agent>
</metsHdr>
```

2. Descriptive Metadata <dmdSec>

Een voorbeeld van beschrijvende metadata volgens Dublin Core die intern is opgenomen in het METS-document:

```
<dmdSec ID="dmd002">
  <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Dublin Core Metadata">
    <xmlData>
      <dc:title>Alice's Adventures in Wonderland</dc:title>
      <dc:creator>Lewis Carroll</dc:creator>
      <dc:date>between 1872 and 1890</dc:date>
      <dc:publisher>McCloughlin Brothers</dc:publisher>
      <dc:type>text</dc:type>
    </xmlData>
  </mdWrap>
</dmdSec>
```

Merk op dat deze metadata een unieke identifier (ID) "dmd002" heeft waarnaar intern in het METS-document gelinkt kan worden via de Structurele map. Dit maakt het mogelijk om specifieke onderdelen van de beschrijvende metadata aan specifieke onderdelen van een digitaal object te koppelen.

3. Administrative Metadata <amdSec>

Het Administrative Metadata <amdSec>-element bestaat uit vier onderdelen: Technical Metadata <techMD>, Intellectual Property Rights Metadata <rightsMD>, Source Metadata <sourceMD>, Digital Provenance Metadata <digiprovMD>. Een versimpeld voorbeeld van technische metadata is:

```
<techMD ID="AMD001">
  <mdWrap MIMETYPE="text/xml" MDTYPE="NISOIMG" LABEL="NISO Img. Data">
    <xmlData>
      <niso:MIMETYPE>image/tiff</niso:MIMETYPE>
      <niso:Compression>LZW</niso:Compression>
      <niso:ScanningAgency>NYU Press</niso:ScanningAgency>
    </xmlData>
  </mdWrap>
</techMD>
```

Ook hier weer een unieke identifier (ID) "AMD001", waarnaar gerefereerd kan worden in de Structurele map of in de File Section. Een bestand in <file>-element bij een <fileGRP> kan naar deze administratieve metadata verwijzen met de betreffend ID:

```
<file ID="FILE001" ADMID="AMD001">
  <Locat LOCTYPE="URL">http://dlib.nyu.edu/press/testimg.tif</Locat>
</file>
```

4. File Section <fileSec>

Een voorbeeld met twee versies van een orale geschiedenis: een transcriptie in TEI-XML en een geluidsbestand in MP3.

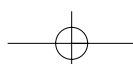
```
<fileSec>
  <fileGrp ID="VERS1">
```

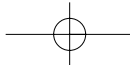
```
<file ID="FILE001" MIMETYPE="application/xml" SIZE="257537" CREATED="2001-06-10">
  <Locat LOCTYPE="URL">http://dlib.nyu.edu/tamwag/beame.xml</Locat>
</file>
</fileGrp>
<fileGrp ID="VERS2">
  <file ID="FILE002" MIMETYPE="audio/mpeg" SIZE="8238866"
    CREATED="2001-05-17">
    <Locat LOCTYPE="URL">http://dlib.nyu.edu/tamwag/beame.mp3</Locat>
  </file>
</fileGrp>
</fileSec>
```

5. Structural Map <structMap>

Het <StructMap>-element geeft de hiërarchische structuur weer in een serie geneste delen en subdelen, aangeduid met <div>. Een <div> kan een mets pointer en file pointer bevatten, waarbij de laatste naar het bestand verwijst. Dit versimpelde voorbeeld representeert een stukje orale geschiedenis dat uit drie delen bestaat.

```
<structMap TYPE="logical">
  <div ID="div1" LABEL="Oral History: Mayor Abraham Beame"
    TYPE="oral history">
    <div ID="div1.1" LABEL="Interviewer Introduction"
      ORDER="1">
      <fptr FILEID="FILE001">
        <area FILEID="FILE001" BEGIN="INTVWBG" END="INTVWND"
          BETYPE="IDREF" />
      </fptr>
      <fptr FILEID="FILE002">
        <area FILEID="FILE002" BEGIN="00:00:00" END="00:01:47"
          BETYPE="TIME" />
      </fptr>
    </div>
    <div ID="div1.2" LABEL="Family History" ORDER="2">
      <fptr FILEID="FILE001">
        <area FILEID="FILE001" BEGIN="FHBG" END="FHND"
          BETYPE="IDREF" />
      </fptr>
      <fptr FILEID="FILE002">
        <area FILEID="FILE002" BEGIN="00:01:48" END="00:06:17"
          BETYPE="TIME" />
      </fptr>
    </div>
    <div ID="div1.3" LABEL="Introduction to Teachers' Union"
      ORDER="3">
      <fptr FILEID="FILE001">
        <area FILEID="FILE001" BEGIN="TUBG" END="TUND"
          BETYPE="IDREF" />
      </fptr>
      <fptr FILEID="FILE002">
        <area FILEID="FILE002" BEGIN="00:06:18" END="00:10:03"
          BETYPE="TIME" />
      </fptr>
    </div>
  </div>
</structMap>
```





De Editorial Board bestaat uit een groep vrijwilligers die geselecteerd zijn uit de internationale METS-community. Instellingen die vertegenwoordigd zijn in de Editorial Board, zijn onder andere Stanford University, California Digital Library, British Library, University of California, Library of Congress, State and University Library in Göttingen, Bodleian Library in Oxford, University of Illinois (Urbana-Champaign), OCLC, Harvard University en Massachusetts Institute of Technology. METS wordt gesponsord door de Digital Library Federation en een consortium van bibliotheken. De Library of Congress treedt op als onderhoudsagentschap door een website te hosten met documentatie en informatie, en andere ondersteunende diensten te bieden.

Ontwikkeling, status en toepassingen

De eerste versie van METS dateert uit 2001. Versie 1.6, de huidige versie, is in oktober 2006 tot stand gekomen. In 2004 kreeg de standaard een registratie van NISO, de internationale organisatie voor informatiestandaarden; deze registratie werd in 2006 hernieuwd. Via een wiki kan men suggesties doen voor versie 1.7. De Editorial Board komt regelmatig bij elkaar en werkt aan de verdere ontwikkeling van het METS-schema, de opname van externe schema's, registratie van profielen, documentatie en educatie. Het gebruik van deze standaard is continu toegenomen sinds

2001. Het wordt toegepast in universiteitsbibliotheken, archieven en musea. Het totaal aantal implementaties is onbekend. In het implementatieregister zijn circa vijfendertig implementaties of projecten officieel geregistreerd die de standaard toepassen. Daarnaast zijn er zestien profielen officieel geregistreerd. Enkele voorbeelden van METS-implementaties:

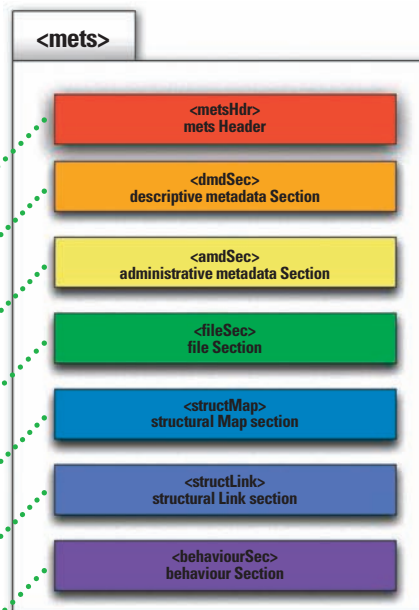
- > OCLC gebruikt de standaard in zijn Digitale Archief service;
- > de Library of Congress past de standaard toe in het Audio-Visual Prototyping Project;
- > de Indiana University heeft een METS Navigator gebouwd om te kunnen bladeren en navigeren in sets van pagina images op het web;
- > RLG (Research Libraries Group) gebruikt de standaard om haar complexe objecten te beheren en op te slaan in de RLG Cultural Materials service;
- > het Göttinger Digitalisierungs-Zentrum gebruikt de standaard in het interne werkproces en voor opslag van digitale objecten.

Renze Brandsma is hoofd van het Digitaal Productiecentrum van de Universiteitsbibliotheek Amsterdam.

Noten

- 1 Voor RDF zie De Standaard, afl. 4, in InformatieProfessional (2006)10, p. 40-43; MPEG DIDL komt in de volgende (en laatste) aflevering aan bod.
- 2 PREMIS is besproken in de vorige aflevering, InformatieProfessional (2007)10, p. 36-38.

Hoofdonderdelen van METS



1. **METS Header** De METS Header bevat beschrijvende metadata over het METS-document zelf, bijvoorbeeld informatie over wie het document heeft gemaakt, wie het heeft aangepast en wanneer het document is gemaakt (te vergelijken met de naam van de catalograaf die een catalogusrecord heeft gemaakt).
2. **Descriptive Metadata** De sectie beschrijvende metadata kan verwijzen naar beschrijvende metadata buiten het METS-document (bijvoorbeeld een MARC-record in een catalogus) of kan interne ingesloten beschrijvende metadata bevatten, of allebei. Meerdere sets van beschrijvende metadata zijn mogelijk, zowel buiten het METS-document (bijvoorbeeld een link naar een record in de OPAC) als intern in het METS-document opgenomen metadata.
3. **Administrative Metadata** De sectie administratieve metadata geeft informatie over hoe de bestanden zijn gemaakt en opgeslagen, intellectueel eigendomsrecht, informatie over de herkomst van de bestanden, etcetera. Net als bij de beschrijvende metadata kunnen de administratieve metadata intern in het METS-document worden opgenomen of via een link naar externe data aan het METS-document worden gekoppeld.
4. **File Section** De sectie bestanden bevat alle bestanden van elektronische versies van het digitale object. <file>-elementen kunnen in een <fileGrp>-elementen worden gegroepeerd, om te zorgen voor een onderverdeling in object-versies.
5. **Structural Map** De structurele map is het hart van een METS-document. Het geeft de hiërarchische of samengestelde structuur van een object uit een digitale bibliotheek aan en linkt de elementen van die structuur aan bestanden en metadata die tot dat element behoren.
6. **Structural Links** De sectie structurele links van METS biedt de mogelijkheid om hyperlinks tussen knopen in de hiërarchie van de structurele map op te slaan. Dit is met name van belang om websites te archiveren met gebruik van METS.
7. **Behavior** De sectie 'gedrag' kan gebruikt worden om digitale content te linken aan bepaalde softwareprogramma's. <

Verder lezen

Een goede plaats om meer informatie over METS te vinden, is de officiële website bij de Library of Congress: www.loc.gov/standards/mets. Ook de implementatie-registry, waaruit de voorbeelden gehaald zijn van instellingen die METS gebruiken, is via deze website toegankelijk: www.loc.gov/standards/mets/mets-registry.html.

