Fondly Collisions: Archival hierarchy and the Europeana Data Model

Valentine Charles and Kerstin Arnold

8th October 2014, DCMI2014, Austin
Overview

→ The Archives Portal Europe - Introduction
  • Projects and content
  • Standards and tools

→ Europeana - Introduction
  • Content and ecosystem

→ Europeana - Data models
  • From Europeana Semantic Elements (ESE) to the Europeana Data Model (EDM)
  • Semantic layer
  • Representing hierarchies
Overview

→ An example
  • Representation in EDM
  • How to get there - conversion in the Archives Portal Europe’s dashboard
  • From apeEAD to EDM - a comparison
→ Remaining challenges
→ Preview
The Archives Portal Europe

The projects behind

- APEnet (2009-2012) with 13 partners (incl. Europeana), extended to 19 partners until the end of the project
- APEX (2012-2015) with 29 partners, extended to currently 33 partners

After the projects: Archives Portal Europe Foundation to maintain and enhance the portal, its network and its position within the European (digital) cultural heritage environment
The Archives Portal Europe

→ The current status

- 30 countries represented, 27 providing content
- 683 institutions, 435 providing content
- more than 50M descriptive units in 335K+ EAD files, 2M linked with digital archival objects
- 1044 descriptions of records creators
The Archives Portal Europe

The use of standards - profiles for central use

- apeEAD for finding aids, holdings guides and source guides
- apeEAC-CPF (Encoded Archival Context - Corporate Bodies, Persons, and Families) for descriptions of records creators
- EAG 2012 (Encoded Archival Guide) - for descriptions of archival institutions and their services
- future addition: apeMETS (Metadata Encoding and Transmission Standard) and apeMETSRights for structured information on digital archival objects
The Archives Portal Europe

→ The use of standards - decentralised responsibility

• local and central tools for conversion and validation of data
• local and central tools for creation/edition of data (especially EAG 2012 and apeEAC-CPF)
• central tool for delivery and processing within the Archives Portal Europe
• central tool for conversion and delivery to Europeana
Europeana.eu, Europe’s cultural heritage portal

32M objects from 2,200 galleries, museums, archives and libraries

Text  Video

Image  Sound

3D
Europeana and its ecosystem

- Libraries
- Archives
- Audiovisual collections
- Thematic collections
- National Aggregators
- Regional Aggregators

Archives Portal Europe
Represent complex data structure: the challenge

→ Initial data model: Europeana Semantic Elements (ESE) and its difficult integration with EAD

→ ESE
  - flat model based on Dublin Core
  - item-centric model
  - no links between cultural heritage objects (CHO) or between objects and contextual entities

→ EAD
  - highly structured model
  - hierarchy containing descriptions of digital and non-digital items/objects
  - nested items/objects with references to contextual entities
A new data framework: Europeana Data Model (EDM)

- A data model that re-uses several existing Semantic Web-based models

- Adopt Semantic Web representation principles (RDF)
  - Re-use and mix different vocabularies together
  - Preserve original data and still allow for interoperability
A data provider submits to Europeana a “bundle” of an object and its digital representation(s)
Connect Archives to the other domains

Create a “semantic layer” on top of cultural heritage objects
Representing hierarchies in EDM

Hierarchies in EDM can be built at the level of:

- **ProvidedCHO**
  - the hierarchy focuses on the description of the semantic structure of the object (fonds, series, subseries, file, item)

- **WebResources (images, scans)**
  - hierarchies of the digitised representations of a ProvidedCHO

- **Contextual resources**
  - hierarchies of “non-information resources” or contextual resources such as edm:Agent, edm:Place, edm:TimeSpan, edm:Concept, edm:Event
An example
An example - representation in EDM

```
edm:WebResource
  rdf:about="http://www.argus.bundesarchiv.de/dy30bmer/index.htm"

edm:hasView

ore:Aggregation
  rdf:about="aggregation_dy30bmer"

edm:aggregatedCHO

edm:ProvidedCHO
  rdf:about="providedCHO_dy30bmer"

dc:title

dc:type

Sekretariat Paul Merker im ZK der SED@de

Fonds
```
An example - representation in EDM
An example - Converting a finding aid to EDM
An example – Converting a finding aid to EDM

Choose a conversion type: Full conversion ☑ Minimal conversion

Name of the data provider*:

Type:

Inherent language of the material from higher levels:

Language(s) of the material:

Specify licence:

Specify European rights statements*:

Specify additional rights information:

[Convert] [Cancel]
An example - Converting a finding aid to EDM

Choose a conversion type: 
- Full conversion
- Minimal conversion

Name of the data provider:*

Type:

Inherent language of the material from higher levels:

Language(s) of the material:

Specify licence:

Specify European rights statements:

Specify additional rights information:

Convert Cancel
An example - Converting a finding aid to EDM
An example - Converting a finding aid to EDM

Choose a conversion type: *Full conversion* ○ *Minimal conversion*

- Text
- Take from file (<dao@xlink:role>) if existing

Specify licence:

- Creative Commons ○ Creative Commons CCO ○ Creative Commons Public Domain Mark ○ European rights statements ○ Out of copyright - non-commercial re-use

Specify additional rights information:

Convert Cancel
An example – Converting a finding aid to EDM

Choose a conversion type: Full conversion ✔Minimal conversion

Name of the data provider:

Type:

Inherent language of the material from higher levels:

Language(s) of the material:

Specify licence:

Specify European rights statement:

Specify additional rights information:

Convert  Cancel
An example – Converting a finding aid to EDM

Choose a conversion type: 
- Full conversion
- Minimal conversion

Name of the data provider:

Type:
- Stiftung Archiv der Parteien und Massenorganisationen
- Take from file (repository) if existing
- Text
- Take from file (<dao@linkrole>) if existing

Language(s) of the material:
- Arabic
- Akan
- Ateso
- Take from file (<language@langcode>) if existing
- English
- Ewe
- Fula
- French
- German
- Italian
- Portuguese
- Spanish
- Swahili
- Somali

Specify licence:
- Creative Commons
- Creative Commons CC0
- Creative Commons Public Domain Mark
- European rights statements
- Out of copyright – non-commercial re-use

Specify additional rights information:
- Free access - no re-use
- Orphan work
- Paid access - no re-use

Convert

[Download for more information: European Rights Guidelines]
An example – Converting a finding aid to EDM
An example – Converting a finding aid to EDM

→ quick note on future extensions

• include apeEAC-CPF for conversion to edm:Agent
• use apeMETS/apeMETSRights for more detailed input on digital archival objects, their relations and their rights status
An example – Converting a finding aid to EDM
An example - apeEAD / fonds

<eadheader
  countryencoding="iso3166-1" dateencoding="iso8601" langencoding="iso639-2b"
  repositoryencoding="iso15511"
  scriptencoding="iso15924"
  relatedencoding="MARC21">
  <eadid
    identifier="DE-1958 dy30bmer" countryCode="DE" mainagencycode="DE-1958"
    url="http://www.argus.bundesarchiv.de/dy30bmer/index.htm">dy30bmer</eadid>
  <filedesc> [11 lines]
  </filedesc>
  <profiledesc> [2 lines]
  <revisiondesc>
    <change>
      <date/>
      <item>Converted_apeEAD_version_1.6.0.2</item>
    </change>
  </revisiondesc>
</eadheader>

<archdesc
  level="fonds"
  type="inventory"
  encodinganalog="3.1.4"
  relatedencoding="ISAD(G)v2">
  <did>
    <unitid
      encodinganalog="3.1.1"
      type="call number">DY 30</unitid>
    <unititle
      encodinganalog="3.1.2">Sekretariat Paul Merker im ZK der SED</unititle>
    <unitdate
      calendar="gregorian"
      era="ce"
      normal="1945/1950"
      encodinganalog="3.1.3">1945 - 1950</unitdate>
    <repository>Stiftung Archiv der Parteien und Massenorganisationen der DDR im Bundesarchiv</repository>
  </did>
  <scopecontent
    encodinganalog="summary"> [3 lines]
  <scopecontent
    encodinganalog="summary"> [11 lines]
An example - EDM / fonds

<ore:Aggregation rdf:about="aggregation_dy30bmer">
  <edm:aggregatedCHO rdf:resource="providedCHO_dy30bmer"/>
  <edm:dataProvider>Stiftung Archiv der Parteien und Massenorganisationen der DDR im Bundesarchiv</edm:dataProvider>
  <edm:isShownAt rdf:resource="http://www.argus.bundesarchiv.de(dy30bmer/index.htm)"/>
  <edm:isShownBy rdf:resource="http://www.archivesportaleurope.net/Portal-theme/images/ape/icons/dao_types/europeana/text.png"/>
  <edm:provider>Archives Portal Europe</edm:provider>
  <edm:rights rdf:resource="http://creativecommons.org/publicdomain/zero/1.0/"/>
</ore:Aggregation>

<edm:ProvidedCHO rdf:about="providedCHO_dy30bmer">
  <dc:description>Kurzbeschreibung: - Überliefert sind u. a. Materialien der führenden [4 lines]
  <dc:description>Paul Merker (1894-1969), ehemals führender KPD-Funktionär, wurde 1946 nach [27 lines]
  <dc:identifier>DY 30</dc:identifier>
  <dc:language>de</dc:language>
  <dc:subject>Fonds</dc:subject>
  <dc:title>Sekretariat Paul Merker im ZK der SED</dc:title>
  <dc:type>Fonds</dc:type>
  <dcterms:temporal>1945 - 1950</dcterms:temporal>
</edm:ProvidedCHO>

<edm:WebResource rdf:about="http://www.argus.bundesarchiv.de(dy30bmer/index.htm)">
  <dc:description>Fonds</dc:description>
  <edm:rights rdf:resource="http://creativecommons.org/publicdomain/zero/1.0/"/>
</edm:WebResource>
An example - EDM / fonds

<ore:Aggregation rdf:about="aggregation_dy30bmer">
  <edm:aggregatedCHO rdf:resource="providedCHO_dy30bmer"/>
  <edm:dataProvider>Stiftung Archiv der Parteien und Massenorganisationen der DDR im Bundesarchiv</edm:dataProvider>
  <edm:isShownAt rdf:resource="http://www.argus.bundesarchiv.de/dy30bmer/index.htm"/>
  <edm:isShownBy
    rdf:resource="http://www.archivesportaleurope.net/Portal-theme/images/ape/icons/dao_types/europeana/text.png"/>
  <edm:provider>Archives Portal Europe</edm:provider>
  <edm:rights rdf:resource="http://creativecommons.org/publicdomain/zero/1.0/"/>
</ore:Aggregation>

<edm:ProvidedCHO rdf:about="providedCHO_dy30bmer">
  <dc:description>Kurzbeschreibung: - Überliefert sind u. a. Materialien der führenden [4 lines]
  <dc:description> Paul Merker (1894-1969), ehemals führender KPD-Funktionär, wurde 1946 nach [27 lines]
  <dc:identifier>DY 30</dc:identifier>
  <dc:language>de</dc:language>
  <dc:subject>Fonds</dc:subject>
  <dc:title>Sekretariat Paul Merker im ZK der SED</dc:title>
  <dc:type>Fonds</dc:type>
  <dcterms:temporal>1945 - 1950</dcterms:temporal>
  <edm:type>TEXT</edm:type>
</edm:ProvidedCHO>

<edm:WebResource rdf:about="http://www.argus.bundesarchiv.de/dy30bmer/index.htm">
  <dc:description>Fonds</dc:description>
  <edm:rights rdf:resource="http://creativecommons.org/publicdomain/zero/1.0/"/>
</edm:WebResource>
An example - apeEAD / hierarchy
An example - EDM / hierarchy (I)
An example - EDM / hierarchy (II)

[...]
Some challenges remain

- Representing hierarchies with one level missing metadata
  - when a missing ProvidedCHO is an inner node (i.e. it has children) it could imply losing information about the context of the children CHO's.
Some challenges remain

- Representing hierarchies with one level missing a digital representation
Hierarchical data are available via:

- data.europeana.eu, Europeana Linked Data pilot
- Europeana API
- Structured data for search engines (Schema.org markups)

One of the challenges is to make developers understand a specific structure of the data.
Displaying hierarchical objects has been Europeana’s biggest challenge!

- Searching hierarchical objects

- It should be possible
  - to make search queries on the objects contained in a hierarchy
  - to make search queries in all the levels of descriptions of an object
  - to search an object via information on its context
  - to allow a user to find the different objects of a hierarchy
Presentation of the hierarchies

- A user should be able to grasp a complex object as a whole and understand the relationships between the different objects, of which the hierarchy is composed.

- An object should be always presented in context (digital representation and descriptive metadata if possible).
Miniature Representing The Power Of Rome, In ‘The Abingdon Apocalypse’

DESCRIPTION

This Apocalypse manuscript contains the Latin text of the biblical book of Revelation, written on versos, with a commentary in French on the facing rectos, each with a miniature. The manuscript takes its name from two added inscriptions: one states that it was given to the Benedictine abbey of St. Mary, Abingdon, by a bishop of Salisbury; the other records that it was lent by the abbot and convent of Abingdon in 1362 to Joan, wife of King David II of Scotland. The image contains four compositions, each with a king and a bishop: three are in circles (representing cities) to the left, and one larger one is to the right. They represent the kingdoms of Assyria, Persia, Macedonia and Rome; the power of the first three has passed to the fourth.

FURTHER DETAILS

Geographic coverage: Abingdon (more info)
Date: [1270]
Date of creation: c. 1270-1275
Type: Stillimage
This item consists of:

- The Abingdon Apocalypse (contains 87 items)
  - Volume 7 (1 item)
  - Volume 8 (3 items)
    - Miniature Representing The Fall Of The Jews And The Call Of The Gentiles, In 'The Abingdon Apocalypse'
    - Miniature Representing The Power Of Rome, In 'The Abingdon Apocalypse'
    - Unfinished Miniature Of The Last Judgement, In 'The Abingdon Apocalypse'
  - Volume 9 (3 items)

View all items above

View all items below
Thank you!

valentine.charles@europeana.eu
k.arnold@bundesarchiv.de

www.europeana.eu
www.pro.europeana.eu

www.archivesportaleurope.net
www.apex-project.eu

8th October 2014, DCMI2014, Austin