



ATHENA
Access to cultural heritage
networks across Europe

eContentplus



Chris De Loof
Royal Museums of Art and History



Technical and Metadata Standards

Brussels
16th of December 2009



ATHENA

Technical and metadata standards



Scope of use of technical and metadata standards (key concepts)

- Digitizing process (forms of information, such as text, sound, image or voice, are converted into a single binary code)
- Museums, Libraries, Archives and Audio-visual archives
- Standardization
- How to? / Best practice ?
- What is Metadata?
- Interoperability – information exchange



Technical and metadata standards : Sources



- Technical guidelines as described on the website of Minerva EC, ([MINERVA EC - Technical Guidelines](#))
- Technical guidelines of Europeana version 1 (<http://www.version1.europeana.eu>)
- best practices of Project Athena (<http://www.athenaeurope.eu>)
- best practices of EuropeanaLocal (<http://www.europeanalocal.eu>)

In 2009 we launched a survey to multiple content providers all over Europe. We asked about the use of :

- Technical formats
- Meta data schemes
- thesauri and semantic interpretations
- limitations of IPR





ATHENA



Technical Standards



Technical Standards Standard Survey Conclusions



Standard

- a guideline which is consistently used as a rule, guideline or definition
 - makes life easier and augments the reliability of the data
 - is designed through experience and expertise of interested parties, such as the users, the makers, the ruling authorities.
 - is applied to a product, object, process or service
 - additionally: increases the interoperability
-
- een richtlijn die consistent gebruikt wordt als een regel, richtlijn of definitie
 - maakt het leven gemakkelijker en verhoogt de betrouwbaarheid van data
 - wordt ontworpen door ervaring en expertise samen te brengen door geïnteresseerden zoals gebruikers, producenten, regelgevers
 - van toepassing op een product, object, proces of dienstverlening
 - Bijkomende: verhoogt de interoperabiliteit



Technical Standards Basic Advice



Use Open Standards as much as possible

Why?

- Increase accessibility
- Stimulate reusability
- Reduce dependency from propriety sellers
- Reduce cost



Technical Standards : 3 Use Environments



What do we intend to do?

Master use = create digital archives

Service use = use the digitized object

Discovery use = search and interoperability



Technical Standards : 3 Use Environments : Master



Master use = create archives

Service use = create and use metadata

Discovery use = search and interoperability

- Source, creation, archive
- Techniques : photography, scan, OCR, digital recording
- Often done by the organisation who may use the object
- quality : very high
- Purpose : archiving
- Speed of delivery: very slow
- Intellectual Property Rights (IPR) : high risk





Technical Standards : 3 Use Environments : Service



Master use = create archives

Service use = create and use metadata

Discovery use = search and interoperability

- Add value to the digital object and make it accessible
- Techniques : semi automatic processing
- Image quality : reasonable
- Speed of delivery : reasonable
- Purpose : make the content accessible (e.g. your own website, intranet)
- IPR : reasonable



Technical Standards : 3 Use Environments : Discovery



Master use = create archives

Service use = create and use metadata

Discovery use = search and interoperability



- Make your content or a subset explorable (optimize queries)
- Make your content searchable
- Increase interoperability
- Techniques : automatic exchange mechanisms
- Speed of delivery : high (but!)
- Quality images : low (e.g. thumbnails, sample audio or video)
- IPR : low risk
- Purpose : increase awareness and dissemination of your content



Technical Standards : Recommendations



- Text Recommendations
- Image Recommendations
- Audio Recommendations
- Video Recommendations
- Vector Graphics Recommendations
- Virtual Reality Recommendations



Technical Standards : Text Recommendations



Master	Service	Discovery
XML (preferred)	XHTML; HTML (preferred)	Sample text or image scan
PDF; DjVu (alternative)	PDF; DjVu (alternative) ODF; RTF; Word	

(supplementar
y)



Technical Standards : Image Recommendations



	Master	Service	Discovery
File Format	TIFF	JPEG, PNG	JPEG, PNG
Colour Quality	8 bit greyscale 24bit colour	8 bit greyscale 24bit colour	8 bit greyscale 24bit colour
Resolution (dpi)	600 (photographs) 2400 (slides)	150-200	72dpi
Maximum dimension (pixels)	Not applicable	600	100-200



Technical Standards : Audio Recommendations



	Master	Service	Discovery
File Format	Uncompressed (preferred) : WAV; AIFF Compressed (alternative) : MP3; WMA; Real Audio; AU	Compressed (preferred): MP3; RealAudio; WMA Uncompressed (alternative): WAV; AIFF; AU	Relevant image
Quality	24 bit stereo and 48/96 KHZ sample rate	256 Kbps (near CD quality); 160 Kbps (good quality)	



Technical Standards : Video Recommendations



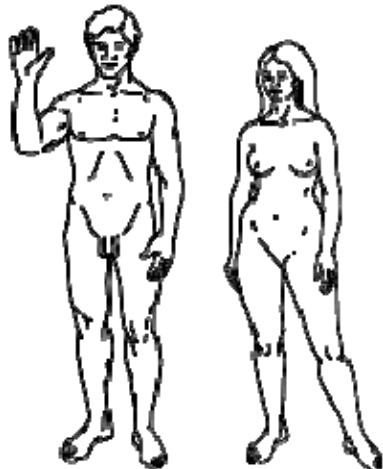
	Master	Service	Discovery
File Format	Uncompressed : RAW AVI (<i>preferred</i>) Compressed (<i>alternative</i>) : MPEG (MPEG1,2 and 4); WMF; ASF; QuickTime	MPEG1; AVI; WMV; QuickTime	Relevant image
Quality	Frame Size : 720x576 pixels Frame rate : 25 frames per second 24 bit colour		

PAL colour encoding



Technical Standards : Vector Graphics Recommendations




Master	Service	Discovery
SVG (preferred) SWF (alternative)	SVG (preferred)	Sample text or image scan 



Technical Standards : Virtual Reality Recommendations



Master	Service	Discovery
X3D (preferred)	X3D (preferred)	Sample text or image scan
Quick time VR (alternative)	Quick time VR (alternative)	



ATHENA



Metadata Standards



Metadata Standards Definition



Traditional definition of Metadata is '**data**' about '**data**'.

Metadata describes other data. It provides information about a certain item's content.

For example, an **image** may include metadata that describes how large the picture is, the colour depth, the image resolution, when the image was created, and other data. A **text document's** metadata may contain information about how long the document is, who the author is, when the document was written, and a short summary of the document.

Possible use of **book** metadata :

- Search for the book,
- Access the book
- Understand about the book before you access OR buy it.



Metadata Standards Basic Advice



Use Standards as much as possible

Why?

- Increase Interoperability and Accessibility
- Stimulate reusability in one or more systems
- Reduce dependency from propriety sellers
- Avoid dependency on a limited number of staff, that is familiar with this system



Metadata Standards : 3 Use Environments



What do we intend to do?

Collections Management : creation of the metadata

Service use = Users get meaningful access to a single piece of metadata describing the object

Discovery use = search and interoperability



ATHENA

Metadata Standards : 3 Use Environments : Collections Mgmt



Collections Management : creation of the metadata

Service use : meaningful access

Discovery use : search and interoperability

Sources of metadata

- Activities and dates related to the management of collections : purchase, loan, reproduction
- Description of the object itself : name, material, dimensions, ...
- Activities and dates related to the life cycle of the object : production, use, restoration, ...
- Activities and dates related to persons, actors, organisations
- Activities and dates related to places, geography



ATHENA

Metadata Standards : 3 Use Environments : Collections Mgmt



Collections Management : creation of the metadata (2)

Service use : meaningful access

Discovery use : search and interoperability

Key Concepts

- Maximum detail
- Goal : Preservation
- Domain : specific (Museums, Libraries, Archives, Audio Visual Archives)
- Speed of delivery: slow, a lot of manual labour



Metadata Standards : 3 Use Environments : Service



Collections Management : creation of the metadata

Service use : meaningful access

Discovery use : search and interoperability

Key Concepts

- Subset of metadata
- Goal : Publication of own content on your website
- Domain : probably specific, probably cross-domain
- Speed of delivery: semi-automated (reasonable)
- IPR : copyright statement



Metadata Standards : 3 Use Environments : Service



Collections Management : creation of the metadata

Service use : meaningful access

Discovery use : search and interoperability

Key Concepts

- Minimal subset of metadata for optimization of queries
- Goal : Maximum spread, indexing and relevance of query results
- Domain : cross domain
- Speed of delivery: Automated (possible fast)
- IPR : reduced copyright

- HARVESTER STANDARDS



ATHENA

Metadata Standards : Best practise Collections management



Domain	Best practice standard
Museum	Spectrum / CDWA
Libraries	MARC
Archives	ISAD(G); EAD



ATHENA

Metadata Standards : Spectrum



Standard for documenting the collection management. Built around 21 procedures, commonly used in museum. Centralised around “units of information”, this is the whole of data which is needed to support the procedures. Developed in the UK, but currently also widely spread in Flanders and the Netherlands.

(English and Dutch descriptions available on the website of Collections Trust: <http://www.collectionstrust.org.uk>).

FARO actively supports the Dutch version of the spectrum in Flanders. (<http://www.faronet.be/blogs/spectrum/download-spectrum>).





ATHENA

Metadata Standards : CDWA



Standard as published by Getty, meant to harmonise the data in different art information systems as much as possible, by adhering the objects and photographs to the conceptual framework. Based on CDWA different compatible subsets can be produced.

http://www.getty.edu/research/conducting_research/standards/cdwa/index.html



Metadata Standards : MARC



Standard for the representation and communication of bibliographic information in machine-readable form.

(<http://www.loc.gov/marc/bibliographic/ecbdhome.html>)





ATHENA

Metadata Standards : ISAD(G) and EAD



EAD (Encoded Archival Description)

W3C Schema used to retrieve and describe electronic archives.

(<http://www.loc.gov/ead>)



Encoded Archival Description _____
_____ Version 2002 Official Site

ISAD(G) General International Standard Archival Description

General rules for archival description that may be applied irrespective of the form or medium of the archival material. The rules accomplish these purposes by identifying and defining twenty-six (26) elements that may be combined to constitute the description of an archival entity.

(<http://www.ica.org>)



Metadata Standards : Service use



No specific metadata standard for publishing content on a website.

Publication is a strategic decision on :

- Which technology (open source, vendors based)
- Which content (all collections, part of the collections)
- Which meta data (subset of metadata)
- Which IPR to use : copyright statement, watermarking, reduced photographs
- Which purpose : intranet, internet, e-commerce site, extranet



Metadata Standards : Discovery use



- A metadata standard for interoperability between systems?
- Dublin Core? Dublin Core Metadata Element Set?
- DC : Description of 15 elements : e.g. Title, Actor, Subject, Provider etc...
- DC : Elements are not mandatory and repeatable
- Nevertheless : famous over the world
- DC : limitations, confusion!
- Europeana uses Europeana Semantic Elements (ESE), this is DC + some new elements (e.g. reference to the digital object on line).



Metadata Standards : Discovery use



- Interoperability between Libraries : use of MARC
- Interoperability between Archives : use of ISAD (G) and EAD
- Interoperability between Museums : local initiatives (Museumdat)

One domain can not use the standard of another!

Rich Cross Domain developments :

1. Light Information for Describing Objects (LIDO)
2. CIDOC CRM
3. Europeana : Europeana Data Model (EDM)



ATHENA

Metadata Standards : LIDO Light Information for Describing Objects



- Standard created as a deliverable by WP3 in the project Athena in co-operation with international experts
- Uses CIDOC Conceptual Reference Model (next slide)
- Based on Museumdat schema and informed by Spectrum
- Full support for Multilinguality
- Aligned to Getty's CDWA Lite Schema
- Has display elements and Indexing elements
- In use in BAM Portal, Germany (<http://www.bam-portal.de>)
- Will be used as 'mapping target' within the project Athena and MIMO
- Official working group on the CIDOC conference in Chile this year
- Vendors meeting for implementation of the standard in software packages





ATHENA

Metadata Standards : CIDOC - CRM



•**CIDOC Conceptual Reference Model (CIDOC - CRM)**. This ISO standard represents a conceptual object-oriented model, which provides meaningful documentation or semantics to cultural heritage and museum through extensive ontological descriptions. This standard is also suitable for the interoperability with other domains, such as the libraries and the archives. It is nevertheless a standard, that requires some study to be able to use it, so it is not that accessible. (<http://cidoc.ics.forth.gr>) It can however be suitable to fulfill the demand for exchange of “rich” metadata between the different domains.





Metadata Standards : Europeana Data Model



- Why
 - to define what information is necessary in order to enable the functionality of Europeana
- What
 - Classes, arranged in a taxonomy
 - Properties, arranged in a taxonomy
 - Constraints: domain/range, cardinality of properties
- Who
 - The Europeana experts
- When
 - July 2010, Danube specs
- Find out for yourself
 - <http://version1.europeana.eu/web/europeana-project/plenary2009docs>





ATHENA



Metadata Mapping



ATHENA

Metadata Mapping



Use of metadata schemes

1. The organization uses an international standard for collection management (e.g. MARC)
2. The organization uses its own schema or vendor specific schema
3. The organization initially uses an international standard for collection management but altered and/or added some elements

MAPPING TO A BEST PRACTICE STANDARD WILL BE NECESSARY IN
CASE 2 and 3



ATHENA

Metadata Mapping Example EAD to ESE/DC



EAD example:

```
<controlaccess>
```

```
  <geogname role="country of coverage" source="tgn">United  
  States</geogname>
```

```
  <geogname role="state of coverage"  
  source="tgn">California</geogname>
```

```
  <geogname role="city of coverage" source="tgn">San  
  Francisco</geogname>
```

```
</controlaccess>
```

Becomes

```
<dcterms:spatial>United States</dc:spatial>
```

```
<dcterms:spatial>California</dc:spatial>
```

```
<dcterms:spatial>San Francisco</dc:spatial>
```

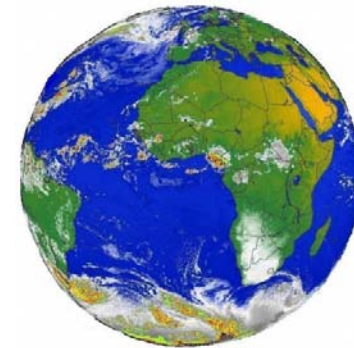


ATHENA

Metadata Mapping In a perfect world...



1. Standard use of metadata schemes for collection management (AED,...)
2. Standard use of thesauri (cfr. workshop thesauri, semantics)
3. Automatic mapping of your metadata schema to the harvester standard of your “choice” (e.g. for Europeana : ESE version 3.2.1) and enrichment with elements from your website (e.g. URL)
4. Optionally semi automatic mapping of your thesauri to controlled vocabularies for the Semantic Web (cfr. workshop thesauri, semantics)
5. Export and automatic delivering of your data (in XML) to your aggregator or Europeana and indexing for search





ATHENA

Metadata Mapping In the real world...



1. No time for mapping
2. No money for digitizing projects
3. Reduced staff
4. Vendor can not export to XML
5. Not strategic enough for management
6. I don't know how to map metadata
7. And my orphan works...
8. Things are changing all the time
9. We are not big enough to do this by ourselves



ATHENA



Conclusions



ATHENA

Conclusions



1. Awareness! You ARE here, this is a good beginning
2. Consider the use of technical and metadata standards
3. Ask for experience (and help others)
4. Read documentation, consider best practices
5. Collaborate! Join projects, in order to deliver digital content to EUROPEANA and to be updated.



ATHENA

Conclusions



Thank you for listening!

Chris De Loof

Project manager IT

Royal Museums for ART and HISTORY (RMAH)

[c.delooof\[a\]kmkg.be](mailto:c.delooof@kmkg.be)

We have some interesting vacancies for the digitization project in the RMAH for 1 year. Interested? Send me your CV 😊.
The ideal opportunity to gain some experience!