Metadata - general introduction

Michael Day

UKOLN, University of Bath
m.day@ukoln.ac.uk

Cataloguing Online Resources: an Introduction to Metadata for Librarians, Manchester, 26 April 2006

http://www.ukoln.ac.uk/

Event timetable

- 09:30 Registration
- **10:00 Metadata - general introduction**
- **10:15 Discovery metadata**
- 11:00 *Break*
- 11:15 Learning Object metadata
- **12:00 Other types of metadata**
- 13:00 *Lunch*
- 14:00 Metadata in practice - JORUM & LOM
- 15:00 Feedback and final discussion
- 15:30 *Close*

Session overview

- Metadata - general overview
  - Definitions
  - Some basic questions
  - Metadata standards

Defining metadata (1)

- Some definitions:
  - Literally, "data about data"
    - Defines the basic concept, but is (perhaps) not very meaningful
    - Refers to everything and nothing (Wendy Duff, 2004)
  - "Machine-understandable information about Web resources or other things" - Tim Berners-Lee, W3C (1997)

Defining metadata (2)

- "Structured data about resources that can be used to help support a wide range of operations - Michael Day, 2001
- "Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage" information objects - NISO, 2004
  - Hints at the many roles metadata can support

Defining metadata (3)

- Metadata is now typically defined by *function*
  - "Data associated with objects which relieves their potential users of having to have full advance knowledge of their existence or characteristics" (Dempsey & Heery, 1998)
  - Popular categorisation:
    - » Descriptive metadata
    - » Structural metadata
    - » Administrative metadata

What functions can be supported?

- Resource disclosure & discovery
- The retrieval and use of resources
- Resource management, including preservation
- Verification of authenticity
- Intellectual property rights management
- Commerce
- Content-rating
- Authentication and authorisation
- Personalisation and localisation of services
- …

To what can metadata be applied?

- "Web resources or other things," e.g.:
- Web sites, Web pages, digital images, databases, books, museum objects, archival records, collections, services, geographical locations, organisations, events, concepts, … even metadata itself

Where can metadata be found?

- Within a resource, e.g.:
  - Title page and table of contents (books), META tags in document headers (Web pages), ID3 metadata (MP3), "file properties" (office documents), EXIF data (images)
- Directly linked to the resource, e.g.:
  - Link rel="meta" elements (Web pages)
- Independently managed in a separate database; can be linked by identifiers
  - This is the most common approach

How important is metadata?

- … "is recognised as a critically important, and yet increasingly problematic and complex concept with relevance for information objects as they move through time and space" -- Gilliland-Swetland (2004)

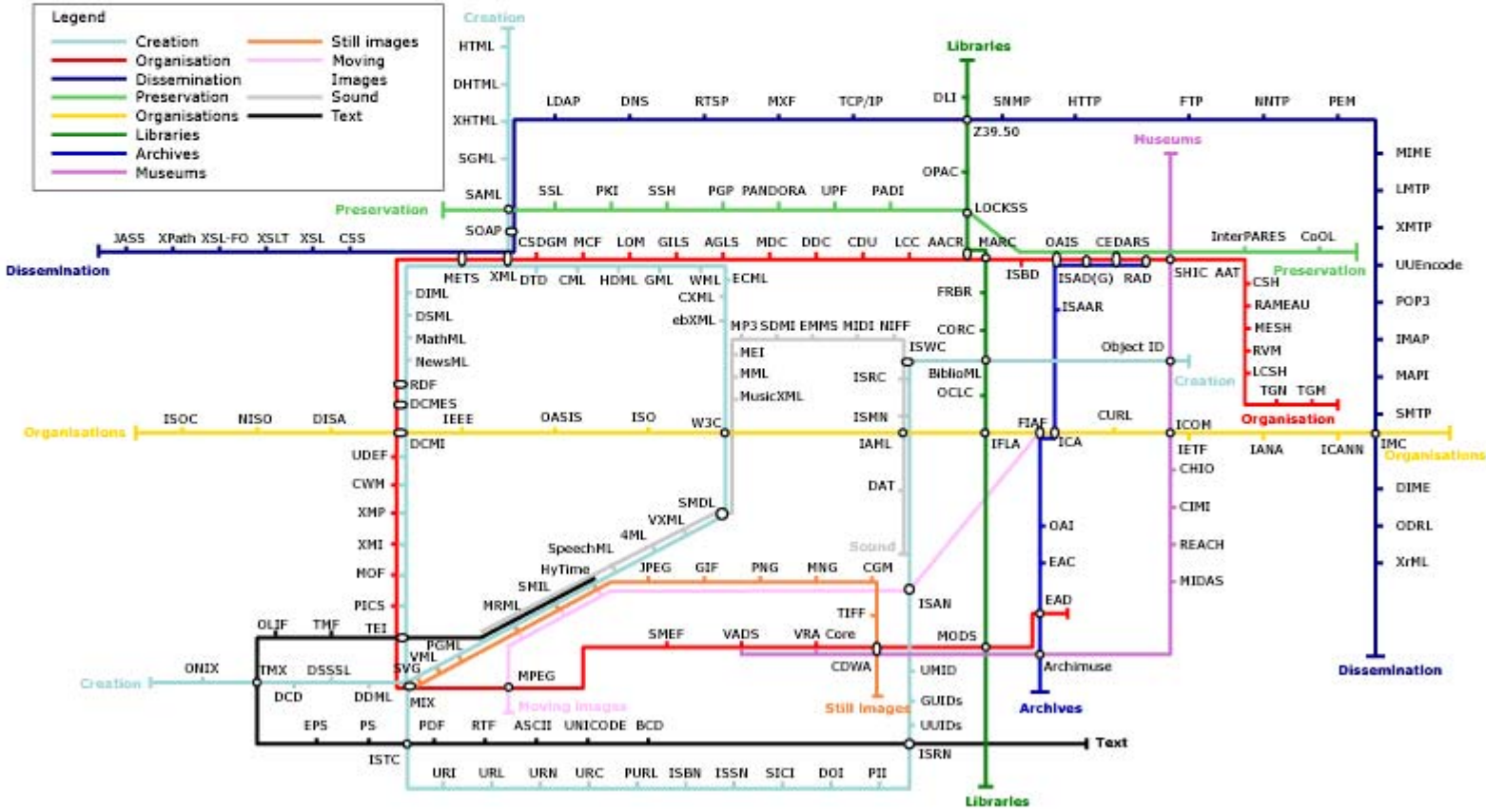
Metadata standards (1)

- But there are a large (and growing) number of metadata initiatives, formats, schemas, etc.
  - See James Turner's MetaMap for one attempt to visualise the metadata information space:
  - **http://mapageweb.umontreal.ca/turner/meta/english/**

MetaMap

[© 2004 James M. Turner, Véronique Moal, & Julie Desnoyers]

Position the mouse over an acronym to see what it stands for in a popup window.
Click on the acronym to see its definition and a link to its official site.



Metadata standards (2)

- Typically defined by "resource management communities"
  - Different traditions, perspectives, functional requirements
- Typically comprise:
  - A "conceptual model" (sometimes not explicit)
  - A set of named components ("terms", "elements" etc) and documentation on their meaning and use
  - A specification of how to represent a metadata instance in a digital format (binding)

Some examples (1)

- Bibliographic:
  - MARC (Machine-Readable Cataloguing) formats, e.g. MARC21
    - Exchange format since 1960s
    - Content often based on family of related standards, e.g. the ISBD series, AACR2
  - MODS (Metadata Object Description Schema)
    - A subset of MARC
  - ONIX
    - Used by publishers and the book trade

Some examples (2)

- Archives and records:
  - ISAD(G) (General International Standard Archival Description)
  - EAD (Encoded Archival Description)
  - EAC (Encoded Archival Context)
  - Recordkeeping metadata (e.g., ERMS (The National Archives), RKMS)
- Museum objects (and collections):
  - SPECTRUM

Some examples (3)

- Digital images:
  - VRA Core, NISO Technical Metadata for Digital Still Images
- Government information:
  - AGLS, e-GMS
- Learning objects:
  - IEEE LOM, UK LOM Core, IMS specifications
- Multimedia:
  - MPEG-7, MPEG-21 (for rights information)

Summing up

- Metadata is ubiquitous
- Metadata enables people and software applications to do things (functions)
  - Not only about "discovery"
  - Different functions require different metadata
- There are many different standards
- Challenges remain in working across standards, or in using standards in combination

Acknowledgements

UKOLN is funded by the Museums, Libraries and Archives Council, the Joint Information Systems Committee (JISC) of the UK higher and further education funding councils, as well as by project funding from the JISC, the European Union, and other sources. UKOLN also receives support from the University of Bath, where it is based.

http://www.ukoln.ac.uk/

The *Digital Curation Centre* is funded by the JISC and the UK Research Councils' e-Science Core Programme.

http://www.dcc.ac.uk/