ECHO Metadata Model

- This model originated from our experience in the ECHO project (European Chronicle On-line)
  - ECHO is an EC funded IST project
  - ECHO aims at providing
    - remote access to collection of historical documentary audio-video resources
    - a software infrastructure to support digital video archives
    - an extensible and interoperable architecture

The approach

Extends the IFLA-FRBR model

Four entities used to describe different aspect of a resource:

- WORK
  - Describes a distinct intellectual or artistic creation
  - Intellectual or artistic realization of a work in the form of alphanumeric, musical, or choreographic notation, sound, image, etc.
  - No information on the physical embodiment is given
  - Examples of WORK are
    - The terrorist attack at Banca Nazionale dell'Agricoltura
    - 2001: A space Odyssey, …

- EXPRESSION
  - Intellectual or artistic realization of a work in the form of alphanumeric, musical, or choreographic notation, sound, image, etc.
  - No information on the physical embodiment is given
  - Examples of EXPRESSION are:
    - TV news on the terrorist attack
    - A documentary on the terrorist attack
    - Interviews on the terrorist attack

- MANIFESTATION
  - Physical embodiment of an expression
  - E.g. manuscripts, books, maps, sound, CD_ROM
  - A single exemplar of a manifestation
  - A documentary on the terrorist attack
  - Interviews on the terrorist attack

- ITEM
  - A single exemplar of a manifestation
  - Versioni
    - Negativo fotografico
    - Copie del negativo
    - Stampe fatte dal negativo
    - Versione digitalizzata del negativo
    - Copie su CD
    - Versione su sito web
**Modello IFLA**

**Work.** Idea astratta di un’opera, per es.,
- L’Iliade, 2001 Odissea nello spazio
- Il sito web del dipartimento di lettere
- La quinta sinfonia di Beethoven
- ecc.

**Expression.** Un work si realizza attraverso un’espressione, per es.
- L’Iliade ha espressioni orali, espressioni scritte, filmati, ecc.
- 2001 odisea nello spazio ha una versione scritta in formato libro, una versione scritta in formato sceneggiatura, una versione filmata, ecc.
- Un’opera musicale ha lo spartito e le esecuzioni
- Il software ha il codice sorgente e l’eseguibile
- Vi sono molti work che hanno una singola espressione.

**Manifestation.** Una expression può essere fornita in una o più manifestations, per es.
- Il testo dell’Iliade è stato fornito in numerosi manoscritti e libri a stampa
- Un’opera musicale può essere distribuita su CD, oppure trasmessa per radio o per televisione

**Item.** Quando si hanno diverse copie di una manifestation, ognuna di esse è un item separato, per es.
- Una specifica copia del libro
- Una specifica rappresentazione dell’opera musicale
- Uno specifico CD
- Ecc.
Ereditarietà. Relazioni di tipo “is a”

- definiscono una relazione di specializzazione di una classe più generale
- Tutti gli attributi e i metodi vengono ereditati dalla classe figlia

Altre relazioni (cont.)

- es.: “composed of”

Altre relazioni (cont.)
Running example

- **Newsreel about the “Landing of the Allied Forces to Normandy”**
  - It is composed of several reportages
  - There are several national versions
    - e.g. Italian and French
  - Each version is available on different supports
    - e.g. VHS tapes, MPEG files
  - There are several copies of the VHS tape with different preservation quality
  - There are several copies of the MPEG file with different access speed

**Cercare, Navigare ed editare metadati di tipo Audio/Video**

**Informatica Umanistica**

**Echo Retrieval Interface**
Informatica Umanistica

Logging

Università di Pisa

The retrieval Interface presentation

Università di Pisa

Metadata of the Selected document

List of documents retrieved

Retrieval Interfaces: search for documents using different techniques
Informatica Umanistica

Retrieval tools interfaces

Università di Pisa

Free text search interface

Università di Pisa
<table>
<thead>
<tr>
<th>Strategies of sorting results</th>
<th>Order of sorting</th>
</tr>
</thead>
<tbody>
<tr>
<td>period of creation time</td>
<td>Restrict the search on a specific archive</td>
</tr>
</tbody>
</table>
Number of results

Informatica Umanistica
Fielded search interface

Università di Pisa

Logical connector among fields
The documents in the collection are classified on the basis of Genres, and Themes, Subthemes, and Thematic keywords. The label of these categories are translated in 5 language (English, Italian, Dutch, French, and German).

A document can be associated to one theme, one subtheme, and one or more thematic keywords.
Informatica Umanistica

Image similarity search

Università di Pisa

Informatica Umanistica

Browsing the document structure

Università di Pisa

Free text to search (any language)

Collection

Theme selection

Browsing the document structure
Informatica Umanistica

The multimedia metadata

Università di Pisa
Informatica Umanistica

The fielded search

Università di Pisa
The (Eurospider) CrossLanguage Retrieval on Transcripts

- The transcript of a video can be in Italian, French, German, or Dutch.
- The CL retrieval allows to search the transcripts using any of these 4 languages, and the English language.

Search for the Italian word “ragazza” (girl) and “ragazzo” (boy) found in transcripts.
Italian word “ragazzi” (boys) found in transcripts for the English word “boy”.

German word “maennchen” (small man) found in transcripts.
The Image Similarity search

Editing metadata with the standard strata-based metadata editor of tecmath

- It is the standard metadata editor provided with the retrieval interface of the tecmath commercial system.
Keyframes

Videosegments (strata)

Audio segments

Scenes detected
Transcript

New stratum name
Write the word “bus” in the scene...

Editing text associated with detected scenes
...and save the document.
They concern to the nature catalogued object. In particular in our model we focused on the time/spatial aspect of the multimedia object, i.e., how the audio/video object are divided in scene, shots, etc. (White Box). E.g., if the data were geographic maps, and the metadata would be points (cities), regions (lakes, rivers).

ECHO Model (cont.)

**Attributes:**
- Transcript
- SpeakerID
- Gender
- SpeakerLanguage
- VideoAbstract
- KeyFrame
- Faces
- Objects
- CameraMovement
- Description

**Attributes:**
- AudioLanguage
- Frequency
- Type
- ...
Browsing/Editing the Bibliographic section

AVDocument

Each entity is associated to a different icon:

- Root of the AVDocument
- Version
- Media
- Storage
Editing the metadata

Unassigned and Assigned fields

Some type of metadata fields do not have null values:

- Boolean (true/false)
- Date
- Closed Lists

Unassigned and Assigned fields

Delete/Add objects

- Deleting an object and its children
  - By selecting an object icon and typing the “Del” button on the keyboard it is possible to delete an object and its children.

- Adding a new child entity
  - It is possible to add a new child object by selecting its parent, clicking on the right button of the mouse, and clicking on “Add child…”. It will appear an input box asking for the title of the new object. Note that, for the Storage entity the title corresponds to the “collocation” field of its metadata.
Create, update and delete

- Create
  - move the cursor of the slider and place it at the start time of the shot you want to create, push the button |
  - Push the button |-. A dialog box will allow you to select the start/end timecodes boundaries of the new segment.
Create, update and delete (2)

- **Modify**
  - Push the button ‘[^’ (‘<’). This will change the end (begin) of the selected segment to the begin (end) of the (preceding) succeeding segment.
  - Push the button ‘^<->’. A dialog box will allow modifying the start/end timecodes boundaries of the selected segment.
  - By selecting the segments by means of the mouse button and pressing shift button, it is possible to select more than one segment. The selected segments are highlighted by means of a thicker line on the sliding bar, and their keyframes are contained in a coloured frame. By pushing the button ‘^>’ all the segmented will be merged together. Note that, the new merged segmented will hierarch the textual metadata of the selected segment that appeared in green.

- **Delete**
  - Select a segment and push the button ‘X’ or the key ‘Del’ of the keyboard.

Updating the keyframes

- Clicking on the button “K” the keyframe of the current video is updated by the frame currently shown.

Creating a new AVDocument

- In order to create a new AVDocument from scratch, click on the menu item File→New. A new AVDocument will be created. Proceed creating the other entities of the AVDocument.

Cross Language features

- Some of the metadata fields of the root of the AVDocument are treated as Cross Language by the editor. The metadata fields involved are:
  - Themes
  - Subthemes
  - Thematic Keywords
Regia Demo

- Bibliographic structure browsing
- Add/Removing children to the tree
- Editing metadata (AVDocument ex.)
- Theme, Subtheme, etc. switching language
- Tips
- Activate an “unassignned” field (ex. EventDate)

Regia Demo (cont.)

- Multimedia metadata structure
  - Video, audio, and transcript segmentations
- Video segmentation tree structure
- Audio segmentation tree structure
- Video segmentation editing/browsing
  - Double representation: keyframe/timeline
  - Tooltip
  - Moving/Playing/Transcript
  - Browsing
  - Editing scenes: merge, remove, extend, add.
  - Textual metadata
- Video jumping from tree-view
- Audio segmentation editing/browsing
- Find

Informatica Umanistica

VideoAnnEx Annotation Tool

DEMO

VideoAnnEx Annotation Tool

- The VideoAnnEx annotation tool assists authors in the task of annotating video sequences with MPEG-7 metadata.
- Each shot in the video sequence can be annotated with static scene descriptions, key object descriptions, event descriptions, and other lexicon sets.
- The annotated descriptions are associated with each video shot and are stored as MPEG-7 descriptions in an output XML file.
- The annotation tool also allows customized lexicons to be created, saved, downloaded, and updated.
Graphical User Interface

**DEMO**

1. Opening an mpeg file: shot detection
2. Video Playback window
3. View Panel:
   - Frames in the shot
   - Shots in the video
4. Shot Annotation:
   - Events
   - Static Scene: background static
   - Objects
   - Keywords

**DEMO**

5. Region Annotation
6. Lexicon editing
7. Categories editing
8. Saving annotation
9. Saving lexicon
10. Saving shot keyframes
11. Load complete example
12. Generated xml files:
   - foo.mpeg
   - foo.mp7.xml
   - foo.lex.xml

Download site

Informatica Umanistica

Osservazioni

Università di Pisa

Segments representation

- Simple mono-dimensional

Segments representation (2)

- Multi-dimensional (strata)

Segments representation (3)

- Hierarchical
Importance of the analysis tools

- More sophisticated software can aid the cataloguers during the editing phase:
  - Intelligent tool of cut detection, which do not produce masses of scenes (tunable: camera movement, fade, etc.)
  - Face and object detection tools are very useful
  - Audio analysis and speech recognition

Model or meta-model?

- Sometimes editing tools allow the user to modify the model
- IBM editor allows to modify the lexicon, or
- the tecmath editor allows to add/remove strata
- There must be a separate tool for maintaining the metadata model
  - Confusion
  - Inconsistency, adaptation of other interface (e.g., the retrieval interface)

Usability

- In this talk we did not talk about the user interface usability
- However this issue is very important in audio/video metadata editing tools
- Silver editor of the Informedia project of the Carnegie Mellon University faces this problem
- This tool is more a video editing tool rather than a metadata editor, however the problem faced are similar

Silver tool user interface
Riferimenti

◆ http://www.ifla.org/

◆ C. Ghilli, M. Guerrini, Introduzione a FRBR, Editrice Bibliografica, 2001
