The CEN Standard for the Identification of Cinematographic Works:
A Brief Overview

June, 2008

Ronny Loewy
Deutsches Filminstitut (DIF)

Why yet another standard?

We already have:

- Library standards (AACR, ISBD, MARC, ...)
- Archival standards (ISAD, EAD, ...)
- Museum standards (Spectrum, CIDOC, ...)
- International Standard Audiovisual Number (ISAN)
- MPEG-7 Multimedia Description Scheme
- Standards for television archives (P/META)
- Visual Resources Association VRA Core
- ... and many more to choose from.

Some Requirements

The simple element set should
- be suitable for basic identification of film works
- be useful for OAI-based infrastructures and as a database browsing format

The extended schema should
- cover most of the complexities of detailed catalogue records
- be designed for interoperability with current metadata specifications
- be useful as common external representation for existing databases and as reference for system design

A Two-Part Standard

Part 1 (EN 15744)
- Defines names and semantics of fifteen basic metadata elements

Part 2 (in preparation as EN15745)
- Extended metadata element set
- Data model covering most FRBR entities and relationships
- Verifiable syntax scheme(s) for external representation and data exchange

CWS Part 1

Not unlike Dublin Core Element Set (ISO 15836) ... and to be used for similar purposes

Defines:

- Title
- Series/serial
- Cast
- Credits
- Production company
- Country of reference
- Original format
- Original length

- Original duration
- Original language
- Year of reference
- Identifier
- Genre
- Relationship
- Source

CWS Part 2

Defines:
- Basic entities (work, expression, manifestation, item, agent, event)
- Basic metadata elements (not identical with part 1)
- Element refinements, sub-elements and element attributes
- Relationship types
- Inheritance between basic entities
- Namespace-based extensibility of elements
- XML schema for syntactic validation
- Dublin Core Application Profile (DCAP)
- Conformance levels and verification criteria

CWS Part 2

Does *not* define:

- Detailed elements at the FRBR item level (i.e. film copies, video files)

- Value lists for data elements and attributes. Where applicable, use of existing vocabularies will be recommended.

- Cataloguing rules (i.e. how to determine the various element values from available sources). This will be handled by the FIAF Cataloguing Commission.

- A complete data model for implementors

Levels of Granularity

Existing databases vary widely in levels of detail. Examples:

- Archive A has only one category of title, archive B has four categories, and archive C has categories plus qualifying attributes such as when and where a title was used.

- Archive A records a single year for the creation date, archive B gives precise production time spans (where known), and archive C additionally records when and where the film has been shot.

Levels of Granularity

Different amounts of detail can be reconciled through levels of granularity. Examples:

- 1962 – year of reference; reference is „unknown"

- 1962 – year of reference; reference is „production"

- 1961-1962 – year of reference; reference is „production"; syntax is time-span

- 1962.02.14-1962.04.12 – „production event"; attribute type is „time-span"; event type is „studio (indoor) shootings"

Note: all examples are compliant

Names vs. Authorities

Some catalogues use authority files, others don't.

A record without authority control may have e.g.
Credits / Screenplay: Person / Name: `John Doe`
                        Person / ID: none

A record with authority control can simply have an added identifier:
Credits / Screenplay: Person / Name: `John Doe`
                        Person / ID: `P093948`

.. or a full copy of the authority record.

Inheritance

Inheritance of properties is implied between several basic entities and for some relationships. Examples:

Work *directed by* Federico Fellini
extends to all versions and editions of the film

Multi-part work *produced by* Ognon Pictures
extends to all individual parts of the work

Extensibility

All metadata elements defined in part 2 will belong to a common XML namespace.

This allows for elements from other standards to be added in various places within the schema.

Examples:

- The *Source* element may be filled with a data structure for bibliographic references.

- The *Colour* element for digital copies may be filled with a data structure from MPEG-7.

Conformance and Verification

Most metadata standards do not (yet) have conformance citeria. CWS will introduce some.

Trust-based conformance
- Criteria for different conformance levels
- Criteria for evaluation
- Can only be verified by intellectual evaluation of data samples

Verifiable conformance
- Formal syntax schema
- Can be verified automatically
- Can only verify syntax, not semantics