

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

September, 2008

Ronny Loewy, Detlev Balzer
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
- 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 EN 15745)

- A comprehensive metadata element set
- A data model based on current reference models
- Verifiable syntax scheme(s) for external representation and data exchange



CWS Part 1 (EN 15744)

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

- 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 (EN 15745)

Defines:

- Basic entities: cinematographic work, expression, manifestation, item, agent, event
- Relationship types
- sub-elements, attributes, and data types
- Namespace-based extensibility of elements
- XML schema for syntactic validation, either stand-alone or via NVDL



CWS Part 2 (EN 15745)

Does not define:

- Detailed elements at the FRBR item level (e.g. for collection management); only a summary of holdings
- 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 and Documentation 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 has additional records of 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 year-span
- 1962-02-14—1962-04-12 – "recording event"; data 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.

- **Screenplay: John Doe**

A record with authority control can simply have an added identifier:

- **Screenplay: [John Doe; ID: P038940]**

... or a full copy of the authority record.

- **Screenplay: [Person record
[ID: P038940;
first name: John;
last name: Doe;
born: 1958-02-12 ...]]**



Aggregation of Information

Merging of filmographic data from different sources can produce confusing results:

Search results (4)

Specify Search »

Title	
HURDES LAS TIERRA SIN PAN	Spain
HURDES, LAS (Tierra sin pan) (Terre sans Pain)	Spain
Las Hurdes	Spain
LAS HURDES O TIERRA SIN PAN	Spain

Is this the same cinematographic work?



Aggregation of Information

Directed by Luis Buñuel; produced in Spain, 1933

Original version, mute

Las Hurdes, tierra sin pan [Spanish, original]

1936 sound version in French

Additional credits to Abel Jacquin (narrator); Darius Milhaud (music) and Johannes Brahms (musical score)

Terre sans pain [French, distribution title]

Las Hurdes o tierra sin pan [Spanish, distribution title]

1936 dubbed version in Spanish

Land without bread [English, translated]

Terra senza pane [Italian, translated]

Las Hurdes - Land ohne Brot [German, translated]

based on



Extensibility

Using namespaces, CWS metadata can co-exist with other metadata standards

```
<cws:AgentRelationship>  
  <cws:Activity>Director of Photography</cws:Activity>  
  <cws:AgentInstance>  
    <mpeg7:Person>  
      <mpeg7:Name>  
        <mpeg7:GivenName>Henri</mpeg7:GivenName>  
        <mpeg7:FamilyName>Alekan</mpeg7:FamilyName>  
      </mpeg7:Name>  
      ... (other person details from an MPEG7 authority file) ...  
    </mpeg7:Person>  
  </cws:AgentInstance>  
</cws:AgentRelationship>
```



Not limited to XML

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
```

```
  xmlns:cws="http://filmstandards.org/schemas/EN15745/1.0#">
```

```
  <rdf:Description  
    about="http://example.org/cinema/works/92837">  
    <cws:CinematographicWork cws:descriptionLevel="m">  
      <cws:Identifier  
cws:scheme="http://example.org/cinema/works/">  
        92837  
      </cws:Identifier>  
      <cws:Identifying title>  
        Die einsame Insel [1920]  
      </cws:Identifying title>
```

```
    ...
```

```
  </cws:CinematographicWork>
```

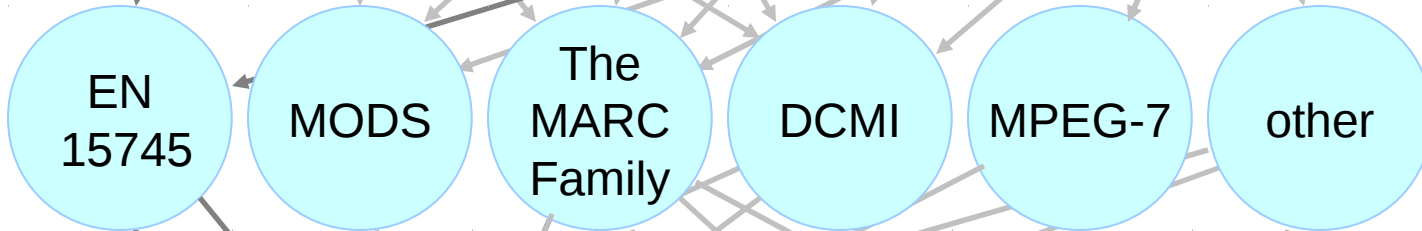


A layer between cataloguing and encoding

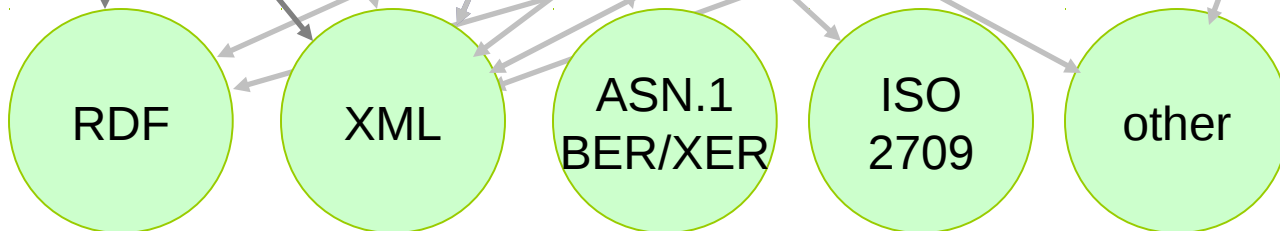
Cataloguing rules



Element sets and structure



Encodings



A part of an infrastructure

