

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 RDF: XML Metadata

8.3 Metadata for Authoring: AAF & SMPTE Standards

8.4 Generic Metadata Framework: MPEG-7

(continued)

Literature:

Rosenblatt/Trippe/Mooney, Digital Rights Management, Chapter 6

Unlabelled Video Tapes & The Internet

- The Unlabelled Video Tape Problem
 - Even worse with digital media: Various formats, variants
- Digital media production:
 - Labelling of parts to be composed
 - » Date, time, format, ...
 - Representing the composition
- Digital media on the Internet
 - Identifying digital media
 - » Title, author, genre, ...
 - Searching for specific media, e.g. audio, video content
 - Fine-grained search within media
 - » e.g. person search within video content
 - Bringing together related media (e.g. text news and photos)
 - » (Automated) syndication

Content, Essence, Metadata

- Content
 - consists of *essence* data and *metadata*
- Essence
 - parts of content that directly represent program material such as audio, video, graphic, still-image, text, or sensor-data
- Metadata
 - parts of content that contain data used
 - » to *describe* essence or
 - » to provide information on its *use*
 - metadata objects sometimes called “mobs”
- Metadata may be
 - Stored separately from the essence data
 - Combined with the essence data (“embedded metadata”)

Source: AAF Developer Overview

Metadata Problems

- Creation metadata
 - During the creation of media essence, metadata is created but often ignored
 - Example: EXIF data in JPEG
- Manually added metadata
 - Users notoriously ignore the administration of metadata
- Metadata incompatibility
 - Metadata exists in various formats specific for media types, applications, product vendors, ...
 - Exchange of metadata is difficult
- Broad range of metadata
 - Metadata exists on various levels, covering all is expensive
- Metadata economy
 - How much of the metadata will be used?
 - When to create metadata?

Selected Metadata Standards

- Dublin Core Metadata Initiative (DCMI) & PRISM (Publishing Requirements for Industry Standard Metadata)
 - Oriented towards books, magazines, journals etc.
 - Uses XML, RDF, Dublin Core
 - www.prismstandard.org
- ONIX (Online Information Exchange)
 - For books: <http://www.editeur.org/8/ONIX>
- TV Anytime (www.tv-anytime.org)
 - Devoted to audio-visual services making use of local mass-storage
 - Focus on: Electronic Program Guide and user profiles
- EBU P/Meta
 - Devoted to material exchange between broadcasting stations
- Commercial solutions by Rovi (www.rovicorp.com), ex Macrovision
 - Company trying to set de-facto industry standard
 - Company collecting large database of media metadata (e.g. www.muze.com)

Dublin Core Metadata Initiative DCMI

- See <http://dublincore.org>
- DC Simple level:
 - Title, Subject, Description, Type, Source, Relation, Coverage, Creator, Publisher, Contributor, Rights, Date, Format, Identifier, Language
- DC Qualified level:
 - Audience, Provenance, RightsHolder, semantic refinements
- Principles:
 - One-to-one: Describes only one manifestation of a resource
 - Dumb-down: Element+value shall be interpretable even when ignoring existing qualifiers

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 RDF: XML Metadata

8.3 Metadata for Authoring: AAF & SMPTE Standards

8.4 Generic Metadata Framework: MPEG-7

(continued)

Literature:

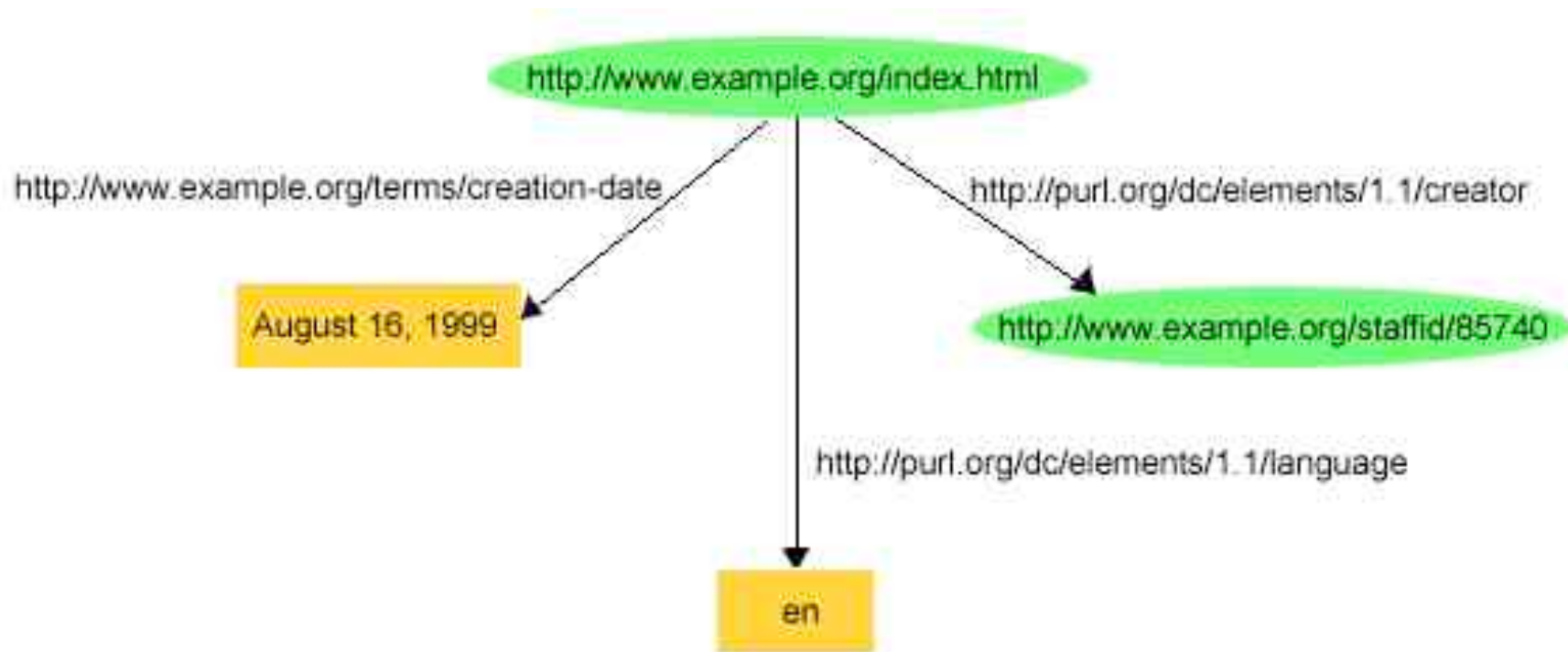
www.w3c.org/RDF

Resource Description Framework RDF

- Language for representing information about resources in the WWW
 - W3C's Semantic Web activity
- *Resource*: Anything that can be identified by a URI (e.g. all Web pages)
- *Property*: An attribute of a described thing which can take on specific values
- *Statement*: A triple consisting of
 - *Subject*: Some resource to be described
 - *Predicate*: A property of the subject
 - *Object*: A specified value
- Properties, values and statements are resources themselves,
 - i.e. can be identified by a URI
 - i.e. can be subject to further description

RDF Example

- `http://www.example.org/index.html` has a **creator** whose value is John Smith
- `http://www.example.org/index.html` has a **creation-date** whose value is August 16, 1999
- `http://www.example.org/index.html` has a **language** whose value is English



RDF/XML Example

- RDF/XML is an XML language for representing descriptions

```
<?xml version="1.0"?>
  <rdf:RDF
    xmlns:rdf=
      "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc=
      "http://purl.org/dc/elements/1.1/"
    xmlns:exterms=
      "http://www.example.org/terms/">
    <rdf:Description
      rdf:about="http://www.example.org/index.html">
      <exterms:creation-date>August 16,1999
        </exterms:creation-date>
      <dc:language>en</dc:language>
      <dc:creator
        rdf:resource="http://www.example.org/staffid/85740"/>
    </rdf:Description>
  </rdf:RDF>
```

Example: Audio Metadata in DC-based RDF/XML

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description
    rdf:about="http://www.medien.ifi.lmu.de/team/
    heinrich.hussmann/files/mmn5a.m4b">
    <dc:creator>Heinrich Hussmann</dc:creator>
    <dc:title>Multimedia Content Description I</dc:title>
    <dc:description>Discusses multimedia metadata
    standards.</dc:description>
    <dc:date>2006-11-24</dc:date>
    <dc:format>audio/mp4</dc:format>
  </rdf:Description>
</rdf:RDF>
```

Example: MusicVocabulary (unofficial!)

- Unofficial proposal for a metadata schema describing classical music and performances
 - <http://www.kanzaki.com/ns/music>
- Defined in OWL-DL (extension of RDF)
- Example:

```
<m:Concert>
  <dc:title>Tokyo Green Symphony Orchestra 12th
    Concert</dc:title>
  <dc:date>2003-11-02T14:00+09:00</dc:date>
  <m:conductorName>Yuri Nitta</m:conductorName>
  <m:performerName>Violin: Tomoko Kawada</m:performerName>
  <m:performerName>Orchestra: Tokyo Green Symphony
    Orchestra</m:performerName>
  <m:programLine>
    Bruch: Violin Concerto No.1 g minor</m:programLine>
  <m:programLine>Brahms: Symphony No.2 D major
  </m:programLine>
  <cal:location>Sumida Triphony</cal:location>
</m:Concert>
```

MusicBrainz RDF Example (deprecated!)

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc  = "http://purl.org/dc/elements/1.1/"
  xmlns:mm  = "http://musicbrainz.org/mm/mm-2.1#">
<mm:Album
  rdf:about="http://musicbrainz.org/album/1d16698f-e1fc-
48b3-ab4a-941c2ecec5bb">
  <dc:title>Rubycon</dc:title>
  <dc:creator rdf:resource="http://musicbrainz.org/artist/
23d8426c-18c7-46e6-a51d-7395bd43c641"/>
  <mm:cdindexidList>
    <rdf:Bag>
      <rdf:li rdf:resource="http://musicbrainz.org/
cdindex/borOdvYNUkc2SF8GrzPepad0H3M-"/>
    </rdf:Bag>
  </mm:cdindexidList>
  <mm:releaseType rdf:resource="http://musicbrainz.org/mm
/mm-2.1#TypeAlbum"/>
  <mm:releaseStatus rdf:resource="http://musicbrainz.org/mm/
mm-2.1#StatusOfficial"/>
  ...
```

Example: Music Metadata Without(!) RDF

- **NO official standard!**
- MusicBrainz XML Metadata Initiative
(<http://musicbrainz.org/doc/MusicBrainzXMLMetaData>)
- Originally: RDF/XML-based standard for audio/video related metadata
 - To be combined with Dublin Core metadata recommendation
- Application: Queries and results for music metadata
 - Similar to CDDB/*Gracenote* but much more flexible
 - Database contents created by community
- Quotation from 2009:
 - "MMD is the official successor of the old RDF-based metadata format, which was popular among semantic web enthusiasts, but didn't have much acceptance otherwise because of its perceived complexity."
 - Using XML-based Web Service for queries to database

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 RDF: XML Metadata

8.3 Metadata for Authoring: AAF & SMPTE Standards

8.4 Generic Metadata Framework: MPEG-7

(continued)

Literature:

AAF Developer Overview, available at www.aafassociation.org

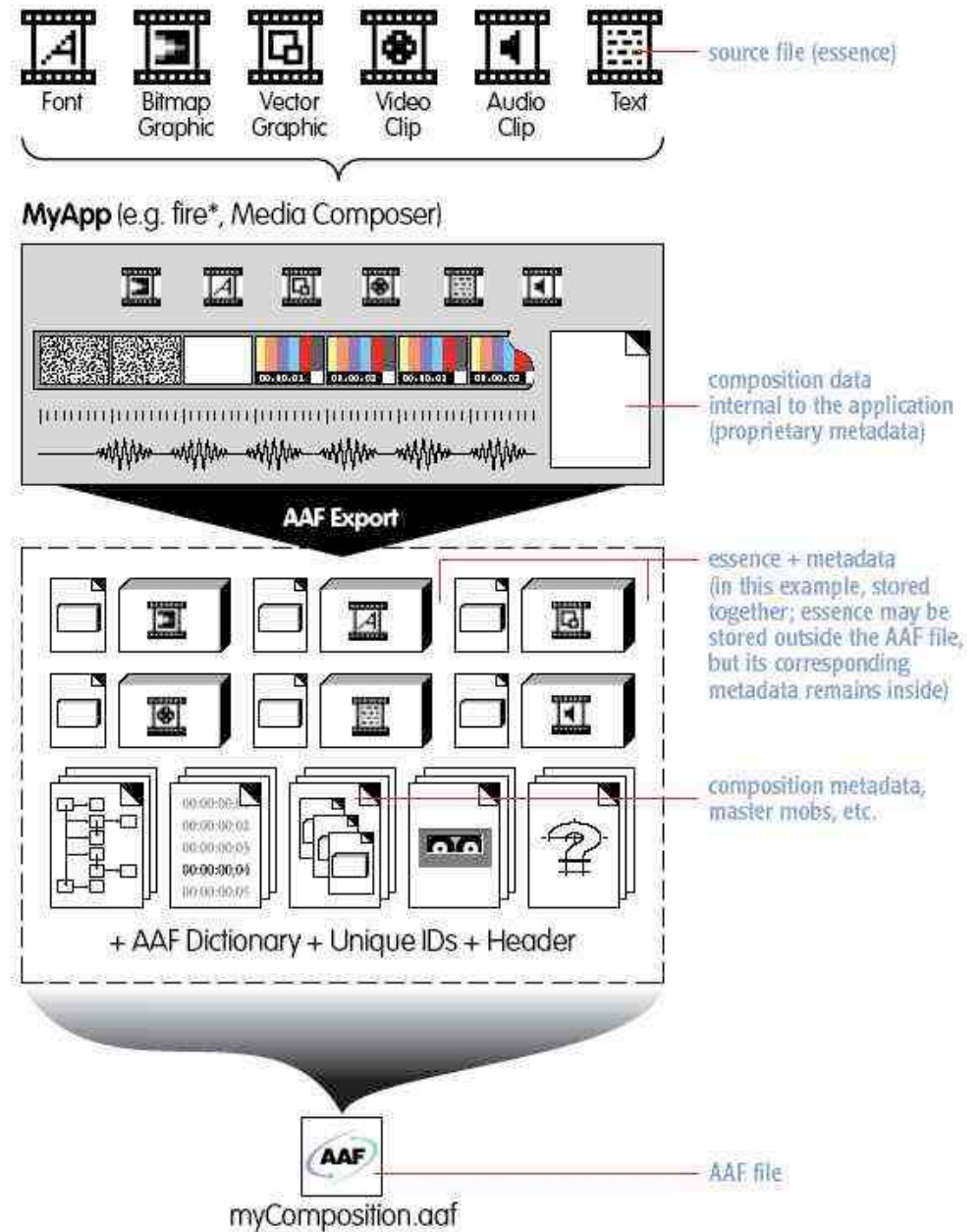
Integration of Digital Media in Video Production

- Example: Putting together all audio elements for a film soundtrack
 - music tracks, ambient sound tracks, performer's synchronized sound, ...
- Multimedia authoring applications
 - Frequently use proprietary data formats
 - Important metadata related to creation process (e.g. compositional metadata) kept only in proprietary formats
- Standards in the broadcasting industry
 - SMPTE (Society of Motion Picture and Television Engineers)
 - EBU (European Broadcasting Union)
 - Working on hardware-based standards for a long time
- EBU/SMPTE Task Force for Harmonized Standards for the Exchange of Program Material as Bit Streams (1996-1999)
 - Results further developed into Advanced Authoring Format (AAF)
 - AAF: Industry-driven, cross-platform, multimedia file format
 - "Advanced Media Workflow Association"

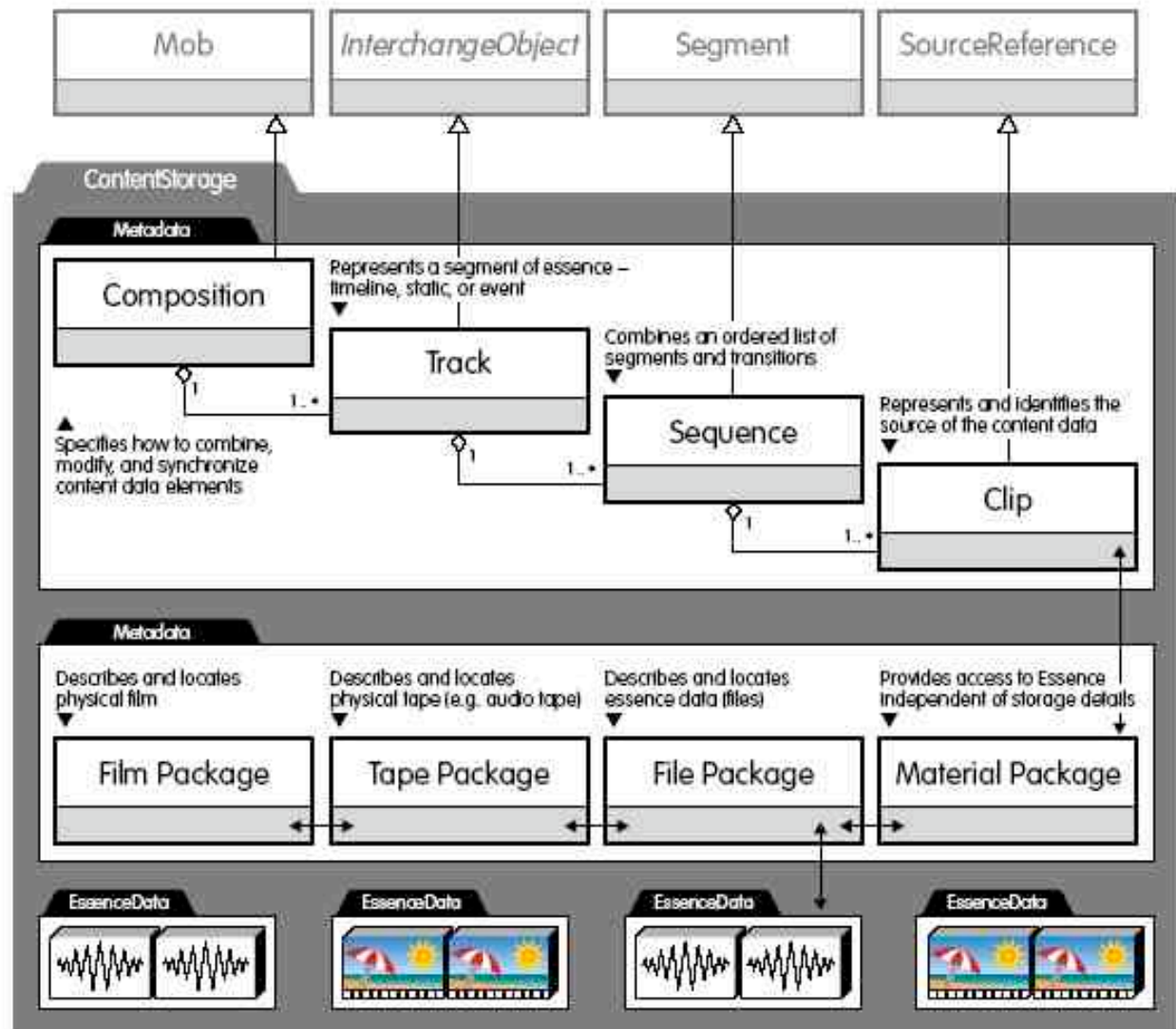
Types of Metadata Covered in AAF

- **Identification and Location Metadata**
 - comprises all forms of metadata that can be used to uniquely identify an item
- **Administration Metadata**
 - definitions of rights, user access, security classifications, encryption, audience listings and other business information.
- **Interpretive Metadata**
 - partly for human-oriented metadata types such as names, artists, organisations and classification.
- **Parametric Metadata**
 - signal coding parameters, device characteristics, sensor parameters (e.g.focal length) plus device storage and streaming parameters
- **Process Metadata**
 - includes all items that describe how essence is assembled, such as editing and compositional metadata
- **Relational Metadata**
 - describes how information is related
- **Spatio-Temporal Metadata**
 - describes places and time including angles, geo-spatial coordinates, dates, creation times, event times, delays and durations

Interchanging Compositions with AAF



AAF Object-Oriented Software Architecture



8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 RDF: XML Metadata

8.3 Metadata for Authoring: AAF & SMPTE Standards

8.4 Generic Metadata Framework: MPEG-7

(continued)

Literature:

www.chiariglione.org

B. S. Manjunath, Philippe Salembier, Thomas Sikora:
Introduction to MPEG-7, Wiley 2002

MPEG-7

- Moving Picture Experts Group (MPEG)
 - = ISO/IEC JTC1/SC29/WG11 “Moving Pictures and Audio”
 - Main Web presence now: www.chiariglione.org/mpeg
- MPEG-7 “Multimedia Content Description Interface”
 - “ ... a standard for describing the multimedia content data that supports some degree of interpretation of the information’s meaning, which can be passed onto, or accessed by, a device or a computer code. MPEG-7 is not aimed at any one application in particular; rather, the elements that MPEG-7 standardizes support as broad a range of applications as possible.”
- Version 1 developed in 1996 – 2001, standard since 2002
- Industrial uptake very slow
 - Ambitious standard
- Some research and open source prototypes available
 - See e.g. <http://www.multimedia-metadata.info>

Parts of the MPEG-7 Standard

- MPEG-7 Systems
- MPEG-7 Description Definition Language (DDL)
 - Descriptors (D) define the syntax and semantics of each *feature* (metadata element)
 - Description schemes (DS) specify syntax and semantics of the relationships between their components, which may be Descriptors or Description Schemes
 - DDL allows the creation of Ds and DSs
 - » XML-based language with some small extensions to XML Schema
- MPEG-7 Visual
- MPEG-7 Audio
- MPEG-7 Multimedia Description Schemes
- MPEG-7 Reference Software
 - eXperimentation Model XM

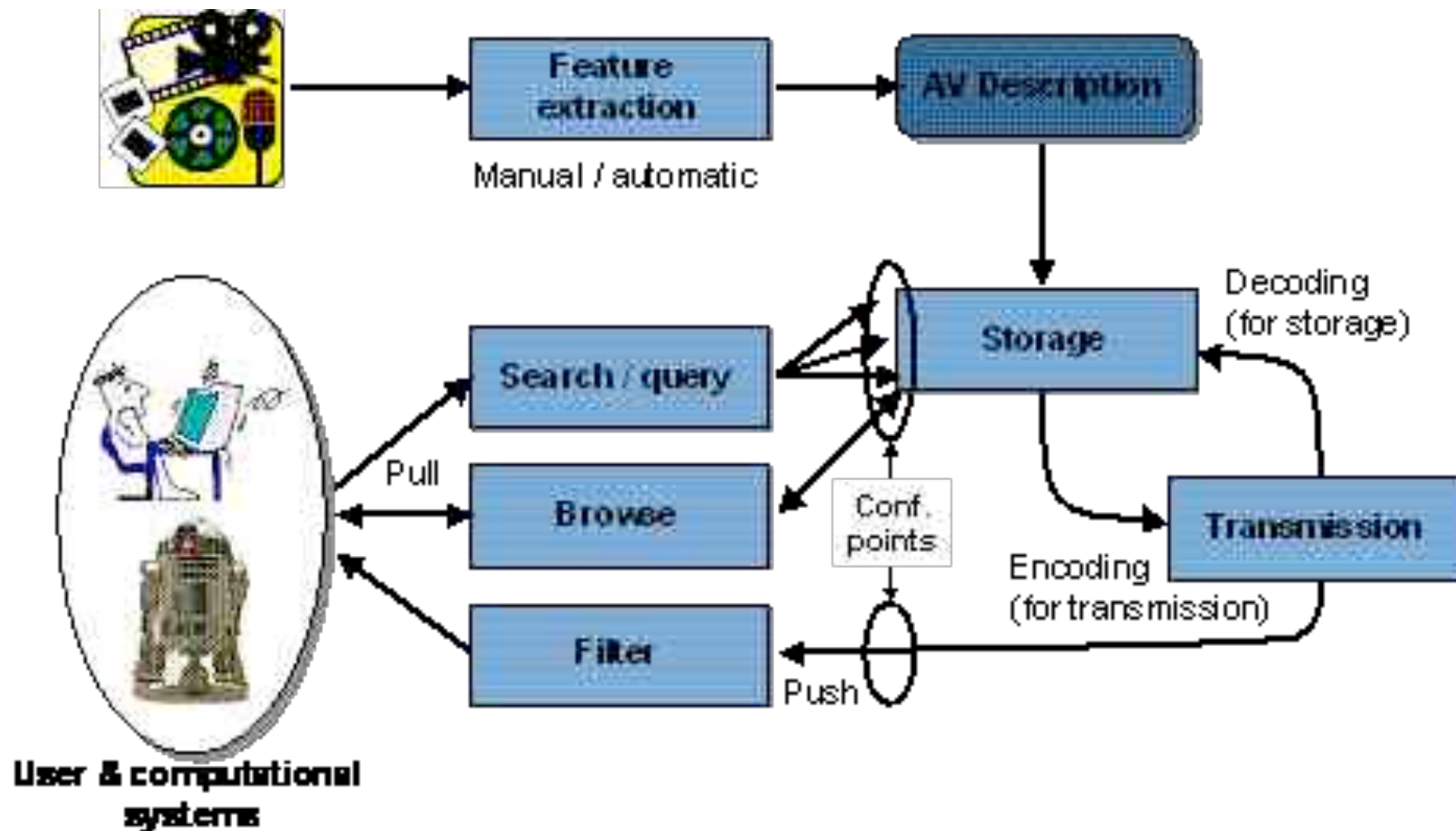
Types of Metadata Covered in MPEG-7

- Technical Metadata:
 - Form (data format, representation parameters like resolution, colour depth...)
 - For live captured material: Time and date of original occurrence
 - Technical parameters of capture (e.g. aperture, exposure etc. for images)
- Content Description Metadata (main focus of MPEG-7):
 - Low level:
 - » Video: Shapes, positions, trajectories etc. of objects (“an object with mainly yellow colour fitting into a box moving from left to right”)
 - » Audio: Key, mood, tempo, tempo changes, position in sound space, ...
 - High level:
 - » Video: “A post car arrives, entering the scene from the left side.”
 - » Audio: Title, composer, structure etc. or, e.g.: “barking dog”
- Additional information:
 - Digital rights, classification, context, further links, ...

Application Areas of MPEG-7

- Architecture, real estate, and interior design (e.g., searching for ideas).
- Broadcast media selection (e.g., radio channel, TV channel).
- Cultural services (e.g., virtual museums).
- Digital libraries (e.g., image catalogue, musical dictionary).
- Education (e.g., repositories of multimedia courses).
- Home Entertainment (e.g., home video management).
- Investigation services (e.g., human characteristics recognition, forensics).
- Journalism (e.g. searching for video footage of political event).
- Multimedia directory services (e.g. yellow pages, Tourist information).
- Multimedia editing (e.g., personalized electronic news service, media authoring).
- Remote sensing (e.g., cartography, ecology, natural resources management).
- Shopping (e.g., searching for clothes that you like).
- Surveillance (e.g., traffic control, surface transportation).
- ...

A Hypothetical MPEG-7 Chain



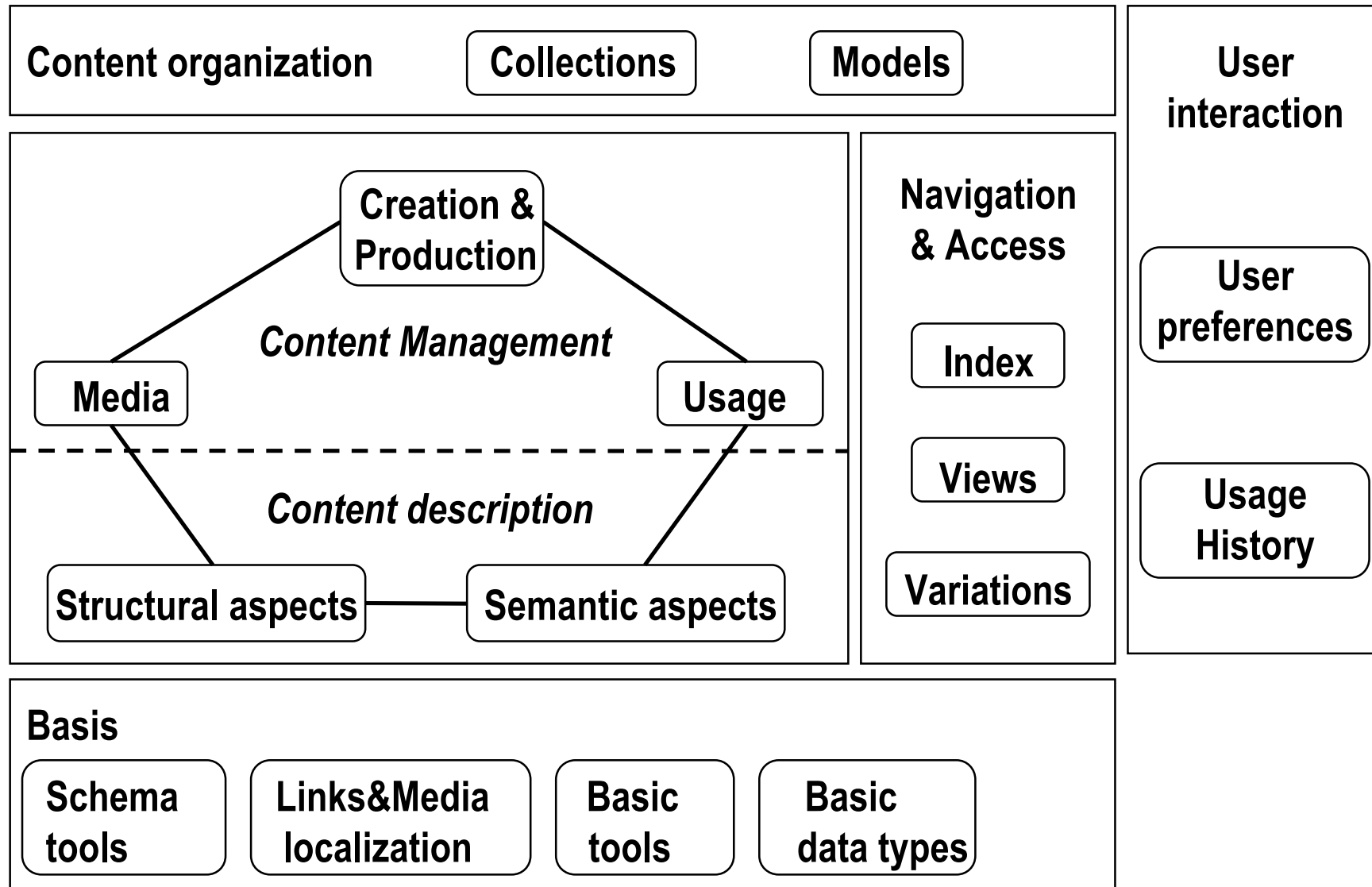
From MPEG-7 Standard Document

Examples of Advanced Queries

- Play a few notes on a keyboard and retrieve a list of musical pieces similar to the required tune, or images matching the notes in a certain way, e.g. in terms of emotions.
- Draw a few lines on a screen and find a set of images containing similar graphics, logos, ideograms,...
- Define objects, including colour patches or textures and retrieve examples among which you select the interesting objects to compose your design.
- On a given set of multimedia objects, describe movements and relations between objects and so search for animations fulfilling the described temporal and spatial relations.
- Describe actions and get a list of scenarios containing such actions.
- Using an excerpt of Pavarotti's voice, obtaining a list of Pavarotti's records, video clips where Pavarotti is singing and photographic material portraying Pavarotti.

From: MPEG-7 Overview

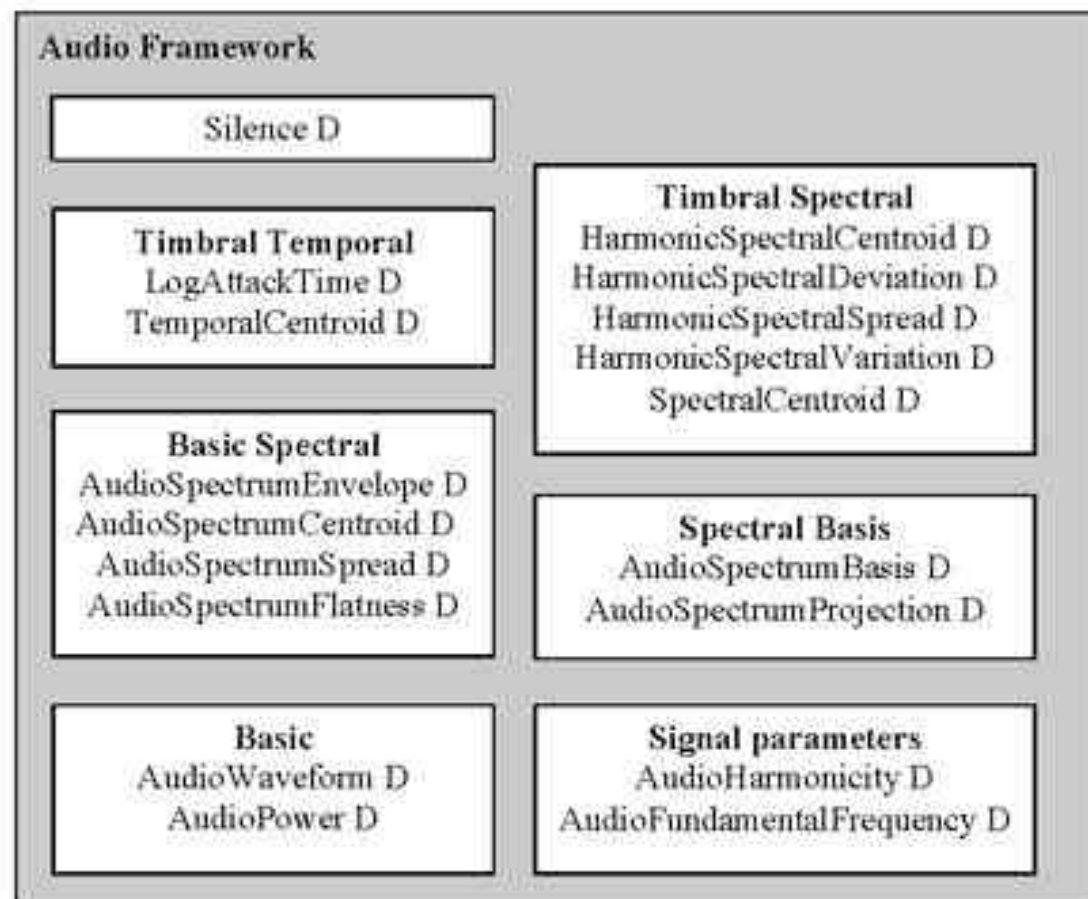
Organization of Multimedia Description Tools



MPEG-7 Audio Low-Level Descriptors

- Structures:
 - Single scalar value
 - Series of sampled values
- Features:
 - See figure

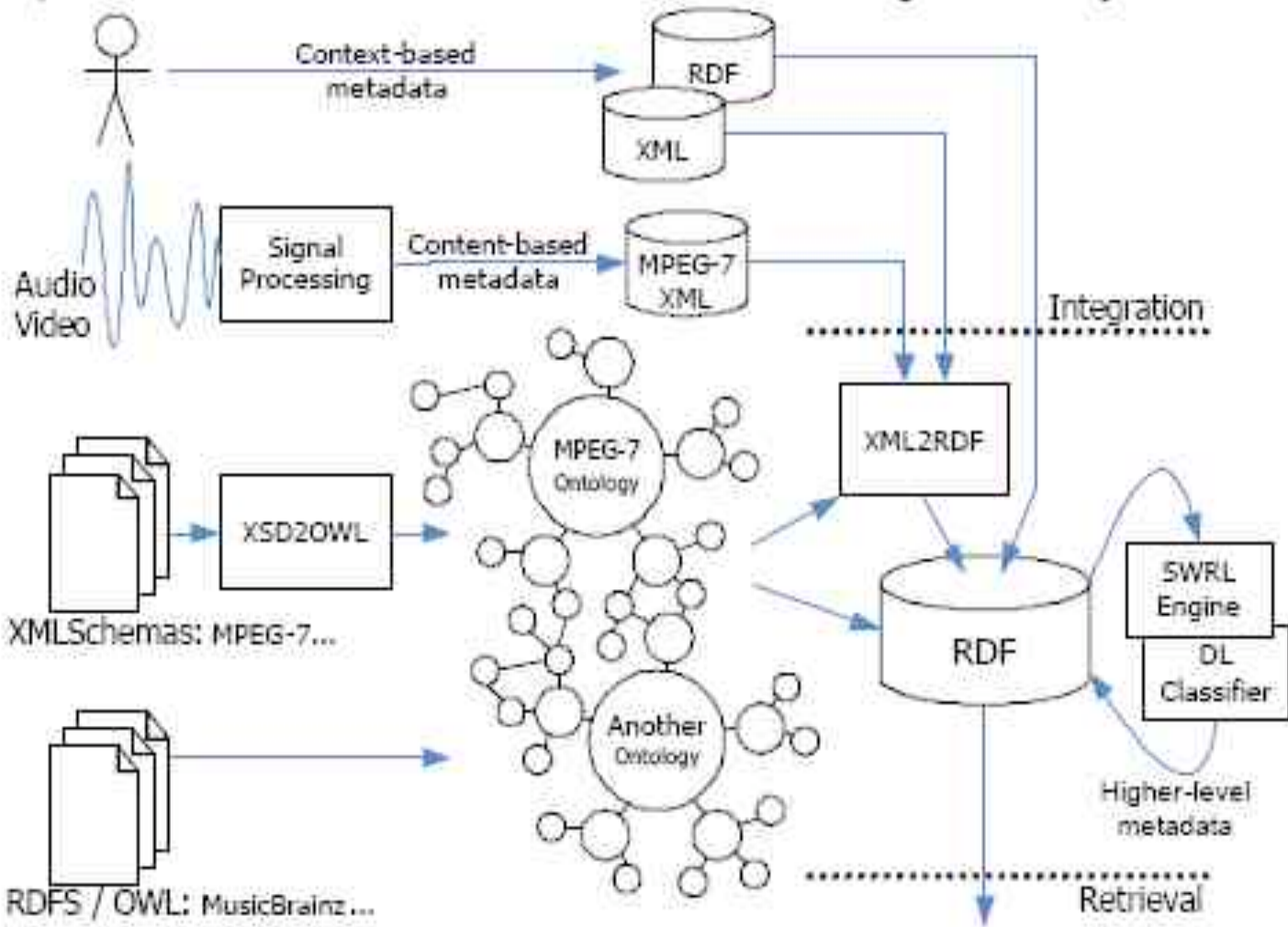
(More details see next lecture)



MPEG-7 Audio High-Level Descriptors

- Audio signature
 - Statistical summary of spectral flatness descriptor
 - Fingerprinting, identification of audio content
- Musical instrument timbre
- Melody description
 - MelodyContour (terse, efficient)
 - MelodySequence
 - » Query by Humming
 - » Example: <http://www.musicline.de/> --> Melodiesuche
- General sound recognition and indexing
 - Probabilistic classifiers for sound classes
- Spoken content
 - Output and intermediate results of Automatic Speech Recognition (ASR)

MPEG-7 Metadata Integration Architecture



Metadata integration and retrieval architecture

Garcia/Selma

Semantic Integration of Different Terminologies

MusicBrainz to MPEG-7 OWL ontology mappings

```
musicbrainz:Artist ⊆ mpeg7:CreatorType
musicbrainz:Album ⊆ mpeg7:CollectionType
musicbrainz:Track ⊆ mpeg7:AudioSegmentType
    dc:author ⊆ mpeg7:Creator
    dc:title ⊆ mpeg7:Title
musicbrainz:sortName ⊆ mpeg7:Name
musicbrainz:duration = mpeg7:MediaDuration
```

Music Vocabulary ontology to MPEG-7 OWL ontology mappings

```
music:Music_Unit ⊆ mpeg7:AudioSegmentType
music:sections = mpeg7:AudioSegment
music:Artist ⊆ mpeg7:CreatorType
```

Garcia/Selma