VK Multimedia Information Systems

Mathias Lux, mlux@itec.uni-klu.ac.at

Dienstags, 16.oo Uhr c.t., E.1.42

Projects

- Kinect-based annotation
- Video quality

Requirements

For projects …

- Prepare a package
  - installation, how to run etc.
- Prepare a presentation
  - 10 min. + 5 min. discussion

Contents

- **Introduction to Metadata**
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

What is Metadata?

Metadata is Data about Data

Meta^2 data is data about metadata

Metadata Applications

- Retrieval & Browsing
  - No need to download / view the whole video
- Management & Organization
  - Rights, Billing, Ordering, Classification
- Adaptation
  - Transformation to appropriate representation
- Service Description
  - Orchestration, Harmonization, Access
  - On technical and semantic level

Metadata Problems

- Interoperability
  - Complexity & power of metadata models
  - Integration in (different) applications & scenarios
- Preservation
  - Readability in 100, 1000 years
  - Description how to decode ...
- Transmission
  - Synchronized, partially, etc.
- Timeliness
  - Changing with audiovisual content while editing?

Aspects of Metadata

- Content Description
- Administrative Aspects
- Quality Metadata
- Legal Metadata
- Technical Metadata

Contents

- Introduction to Metadata
- **Metadata Formats**
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Standards Preface: XML

- eXtensible Markup Language
- Recommendation by the W3C
  - Simplification of SGML
- Base language for many other recommendations
  - SVG, XHTML, SMIL, ...

XML: Overview

- Header identifying version & coding
- Tree-like structure
- Simple structuring elements
  - Tags & attributes (Markup)
  - Entities
- DTD and XML Schema for model definition
  - DTD is ‘simple’ and small
  - XML Schema is XML based and rather powerful

XML Benefits

- Existing Parsers
  - Document Object Model (DOM)
    - Tree structure in memory
    - Access through navigation in tree
  - Simple API for XML (SAX)
    - Event based, sequential
    - Low memory footprint, no random access
- Schema of data has to be defined
  - Parsers implementation is simple

Contents

- Introduction to Metadata
- Metadata Formats
  - **Media Production**
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Media Production: Dublin Core

- Aims to provide
  - Common denominator for metadata
  - Simple yet powerful schema
- Dublin Core Metadata Initiative defined
  - 15 elements (author, date, title, type, ...)
  - Further refinements (creation date, extent, ...)
- Dublin Core does not provide
  - A schema for storage
  - A schema for data types (e.g. dates)

Media Production: EBU P/Meta

- Aims to provide …
  - a universal standard for metadata exchange between professional media organizations
  - a definition of common meaning to the data fields and values that most broadcasters use in order to enable exchange
  - designed for use in a wide range of broadcasting activities
  - both language and system independent
  - a joint development by EBU (European Broadcasting Union) members on a not-for-profit basis
  - a scheme that makes use of other standards where possible, e.g. ISO country codes

Media Production: Other Standards

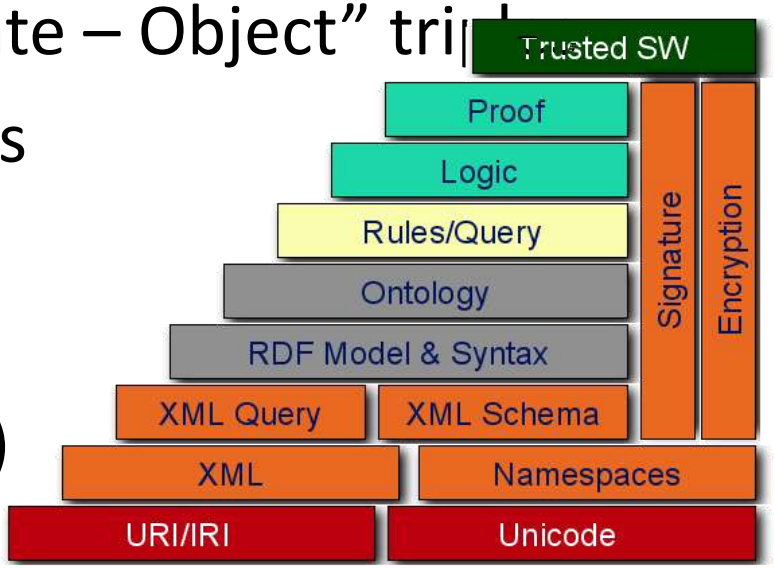
- SMPTE Metadata Dictionary
  - Society of Motion Picture and Television Engineers
    - Since 1916, 61 members
  - Standard for metadata exchange in TV
  - Defines set of attributes / fields
- MXF DMS-1
  - Metadata bundled with the Material Exchange Format (MXF)
  - Open format for the broadcasting area (SMPTE + EBU)
- Virtually ‘no information’ about these is available
  - Just for exchange for insiders
  - Might not be royalty free

Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - **Ontologies**
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Ontologies: RDF

- Metadata Model published by the W3C
  - Reaction on the insufficiency of HTML metadata for search & inference
  - Based on “Subject – Predicate – Object” tri
  - URIs for identifying concepts
  - Spans a directed graph
  - Is used in conjunction with vocabularies (e.g. DC, FOAF)



Ontologies: SKOS

- Simple Knowledge Organization System
  - RDF Vocabulary for KOS
- Knowledge Organization Systems are
  - Taxonomies, Thesaurii, Classification Schemes, etc.
- Can be used to organize multimedia data

Ontologies MMSEM

- Multimedia Semantics : Incubator Activity of the W3C
  - Closed Aug. 2007

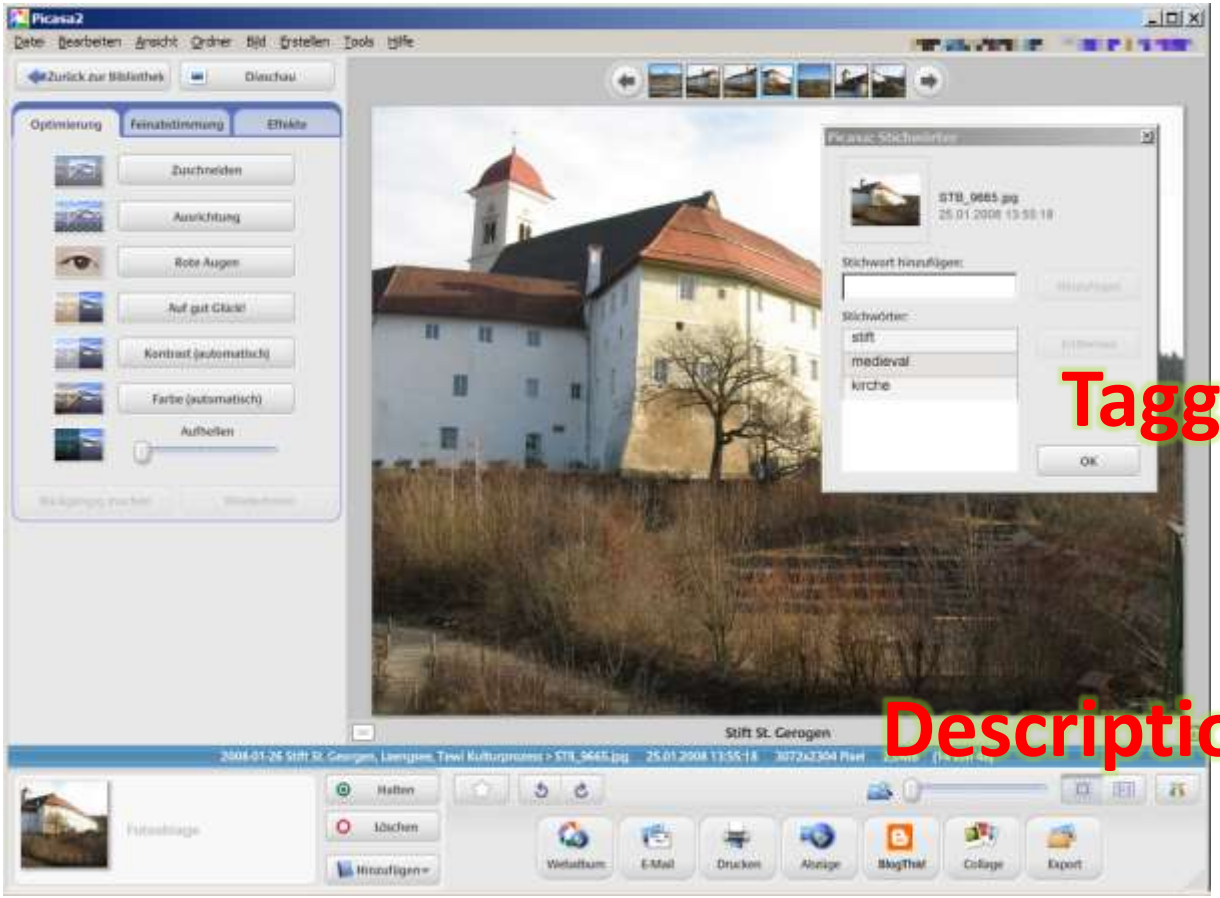
Deliverables:

- Image Annotation on the Semantic Web.
  - use cases and general discussion about Semantic Web vocabularies and tools
- Multimedia Annotation Interoperability Framework.
  - a bottom-up approach to provide a simple extensible framework to improve interoperability
- MPEG-7 and the Semantic Web.
  - four current OWL/RDF proposals of MPEG-7, as well as a comparison of the different modeling approaches in the context of practical applications.

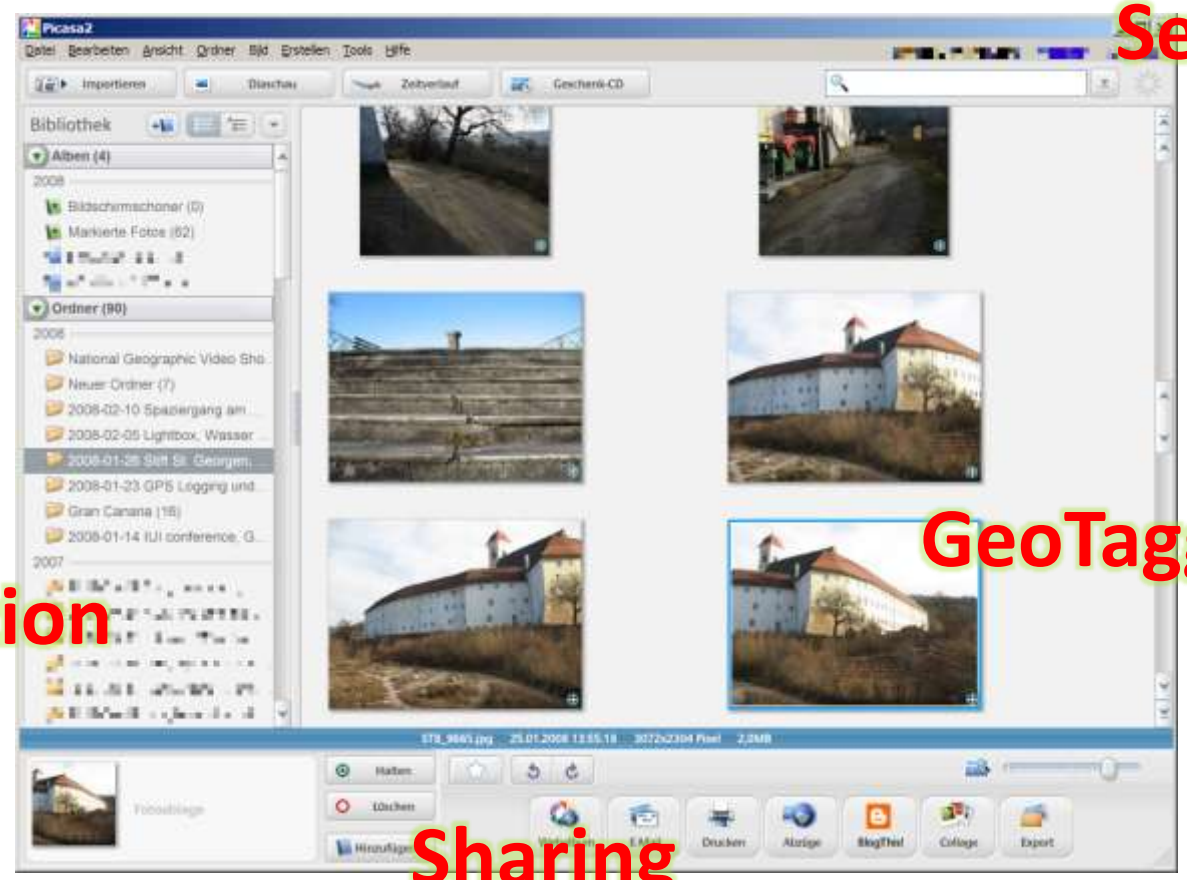
Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - **Home User**
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Home User: Metadata Applications



Home User: Metadata Applications



Home User: EXIF

- Exchangeable Image File Format (EXIF)
  - Japan Electronic and Information Technology Industries Association (JEITA)
  - Extensive format for technical aspects
  - Settings and sensor readings at the time of recording
  - Mostly images from digital cameras

EXIF - Example

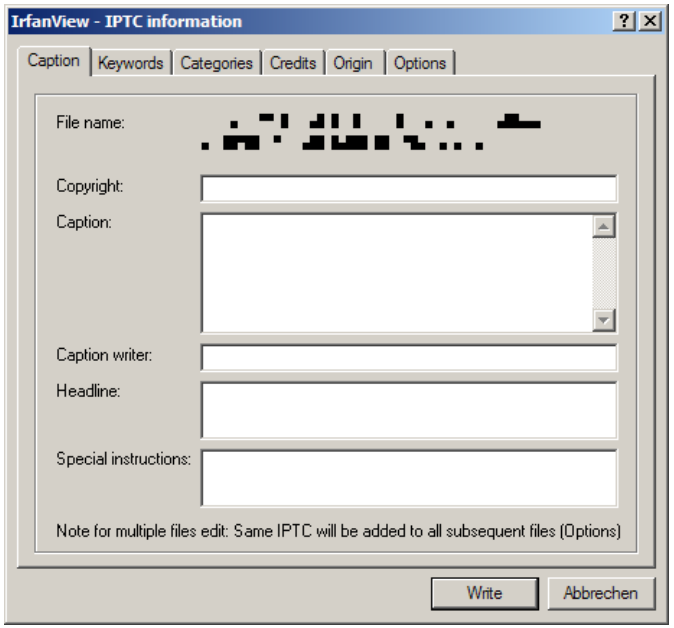
Make - Canon
Model - Canon PowerShot A620
Orientation - Top left
XResolution - 180
YResolution - 180
ResolutionUnit - Inch
DateTime - 2008:02:10 15:44:58
YCbCrPositioning - Centered
ExifOffset - 198
ExposureTime - 1/200 seconds
FNumber - 2.80
ExifVersion - 0220
DateTimeOriginal - 2008:02:10 15:44:58
DateTimeDigitized - 2008:02:10 15:44:58
ComponentsConfiguration - YCbCr
CompressedBitsPerPixel - 5 (bits/pixel)
ShutterSpeedValue - 1/202 seconds
ApertureValue - F 2.80
ExposureBiasValue - 0.00
MaxApertureValue - F 2.80

GPS information: -
GPSVersionID - 2.2.0.0
GPSLatitudeRef - N
GPSLatitude - 46 40 41.41
GPSLongitudeRef - E
GPSLongitude - 13 58 22.17
GPSAltitudeRef - Sea level
GPSAltitude - 503 m
GPSTimeStamp - 14 44 58

Maker Note (Vendor): -
Macro mode - Normal
Self timer - Off
Quality - Superfine
Flash mode - Auto + red-eye reduction
Sequence mode - Single or Timer
Focus mode - Single
Image size - Large
Easy shooting mode - Portrait
Digital zoom - None

Home User: IPTC

- IPTC Information Interchange Model (IIM)
  - Several elements to describe images (assets)
  - Rather common format
    - Adobe Bridge / Photoshop
    - Google Picasa
    - Irfanview …
  - Like a *predefined metadata form* ->



Home User

- eXtensible Metadata Platform (XMP)
  - Initiative from Adobe
  - Based on RDF, embedded in document
  - Also used in PDF, AI, PSD, etc.
- ID3
  - Metadata for MP3, spread by popular players
  - Two versions ...
    - v1: 128 Byte block coding some fields at end of file
    - v2: Several optional tags inside stream

Broadcasting + iTV

- Electronic Program Guide (EPG)
  - In use in conjunction with DVB
  - Simple format in additional stream
- Multimedia Home Platform (MHP)
  - In use in Austrian DVB-T
  - Proprietary format for data + function
  - Based on Java

Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- **MPEG-7**
- MPEG-21
- Metadata Generation & Annotation

MPEG-7

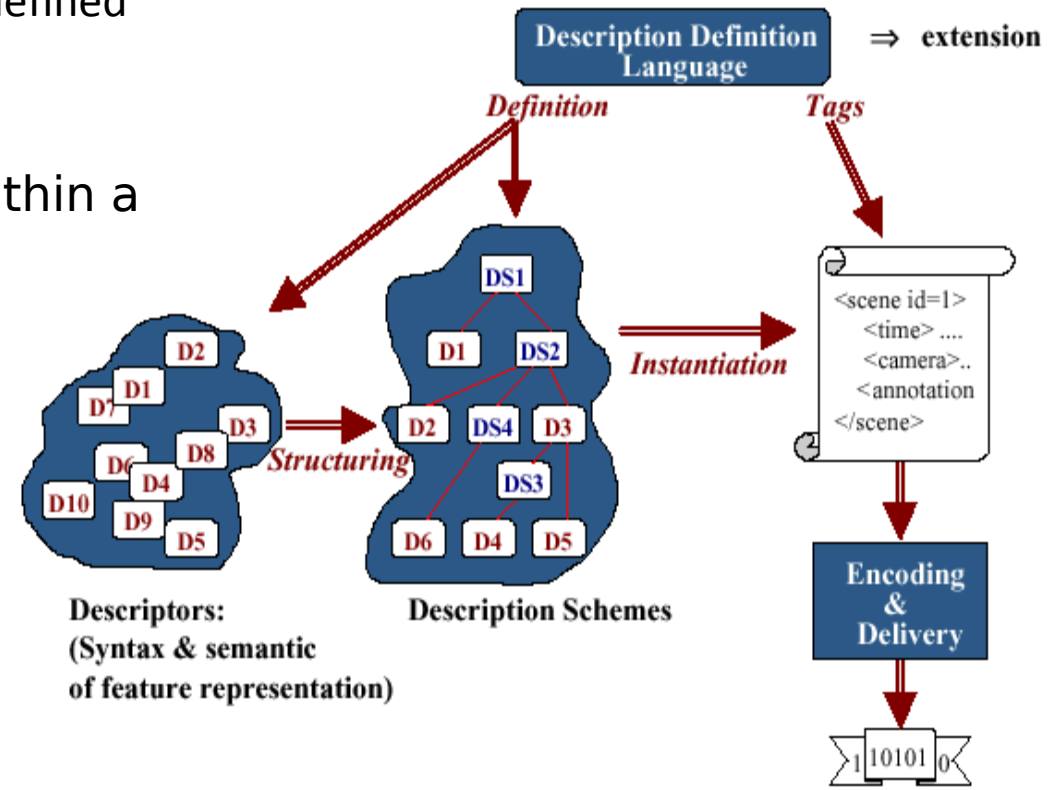
- ISO/IEC Standard: **Multimedia Content Description Interface**
- Moving Pictures Expert Group
  - Specification goes on …
- It's based on XML (Schema)
  - Binary representations possible (BiM)
- Allows differing granularity of descriptions
  - Extensive to very simple

MPEG-7 History

- Call for Proposals: October 1998
- Evaluation: February 1999
- First version of Working Draft (WD): December 1999
- Committee Draft (CD): October 2000
- Final Committee Draft (FCD): February 2001
- Final Draft International Standard (FDIS): July 2001
- International Standard (IS): September 2001

MPEG-7 Basics

- Descriptors
  - Syntax and semantics of exactly one (low or high level) elementary feature
  - Also base data types are defined
- Description Schemes
  - Defines structures within a framework
- Description Definition Language (DDL)
  - Extension of XML Schemes
- Coding Schemes
  - Create and interpret descriptions in BiM



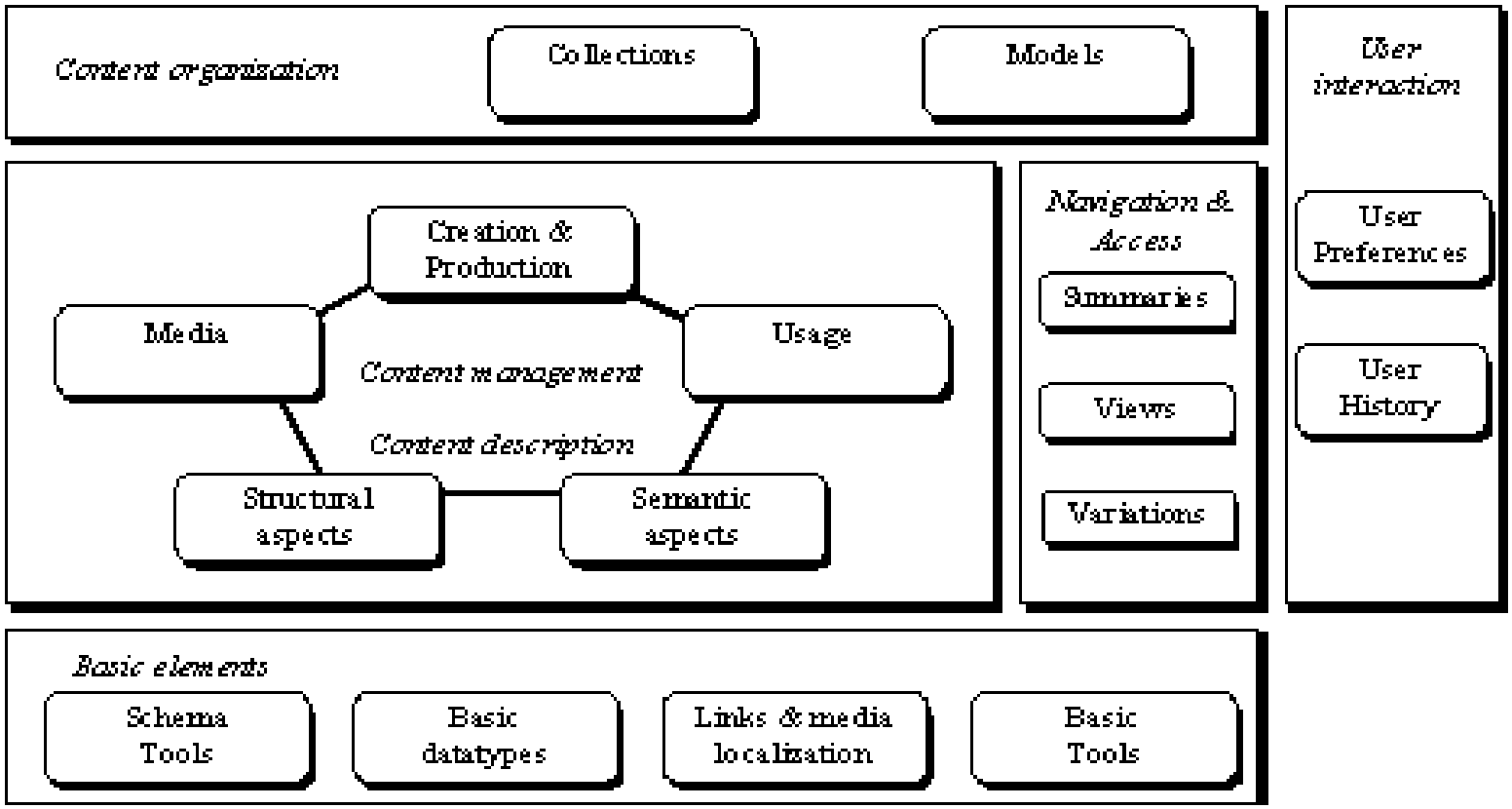
MPEG-7 Parts

1. MPEG-7 Systems
   - Tools needed to prepare MPEG-7 descriptions for efficient transport and storage and the terminal architecture.
2. MPEG-7 Description Definition Language
   - Language for defining the syntax of the MPEG-7 Description Tools and for defining new Description Schemes.
3. MPEG-7 Visual
   - Description Tools dealing with (only) visual descriptions.
4. MPEG-7 Audio
   - Description Tools dealing with (only) audio descriptions.
5. MPEG-7 Multimedia Description Schemes
   - Description Tools dealing with generic features and multimedia descriptions.

MPEG-7 Parts

6. MPEG-7 Reference Software
   - Implementation of relevant parts of the MPEG-7 Standard with normative status.
7. MPEG-7 Conformance Testing
   - Guidelines and procedures for testing conformance of MPEG-7 implementations
8. MPEG-7 Extraction and Use of Descriptions
   - Informative material about the extraction and use of some of the Description Tools.
9. MPEG-7 Profiles and levels
   - Provides guidelines and standard profiles.
10. MPEG-7 Schema Definition
    - Specifies the schema using the Description Definition Language

Scope of MPEG-7



from: http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

Basic Elements

Basic elements are fundamental constructs and used throughout the whole MPEG-7 description

- Basic datatypes
  - Time and date, relative and absolute
  - Numeric datatypes like matrices and vectors
- Links & Media Localization
  - Interconnections and content linking

Navigation & Access

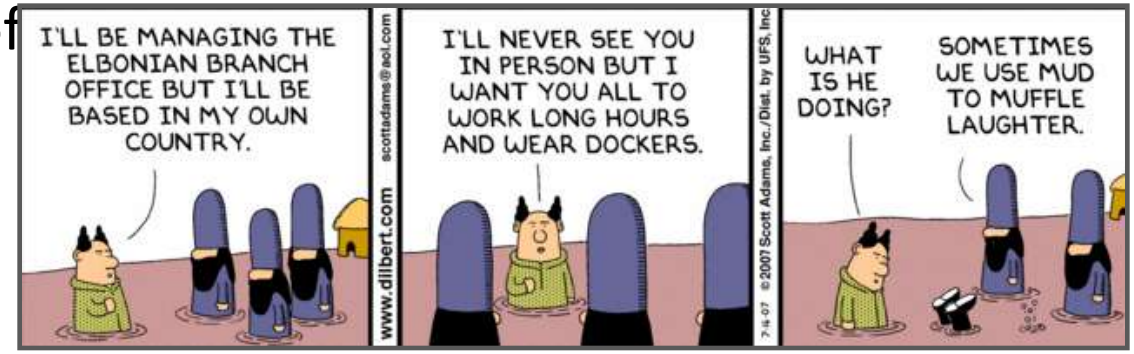
- Descriptors for Browsing & Retrieval
  - Summaries
  - Partitions (time, space & frequency)
  - Decompositions (time, space & frequency)
  - Variations

User Interaction

- Pertaining consumption of AV data
  - user preferences
  - usage history
- Meant to facilitate personalization
  - Matching User Interaction DS with content description
  - Is research topic @ ITEC

Content Organization

- Organization & modelling of collections
  - Audio-visual content, segments, events, and/or objects
    - E.g. pictures, scenes, music files, etc.
  - Allows collection description as a whole
    - E.g. “Pictures of

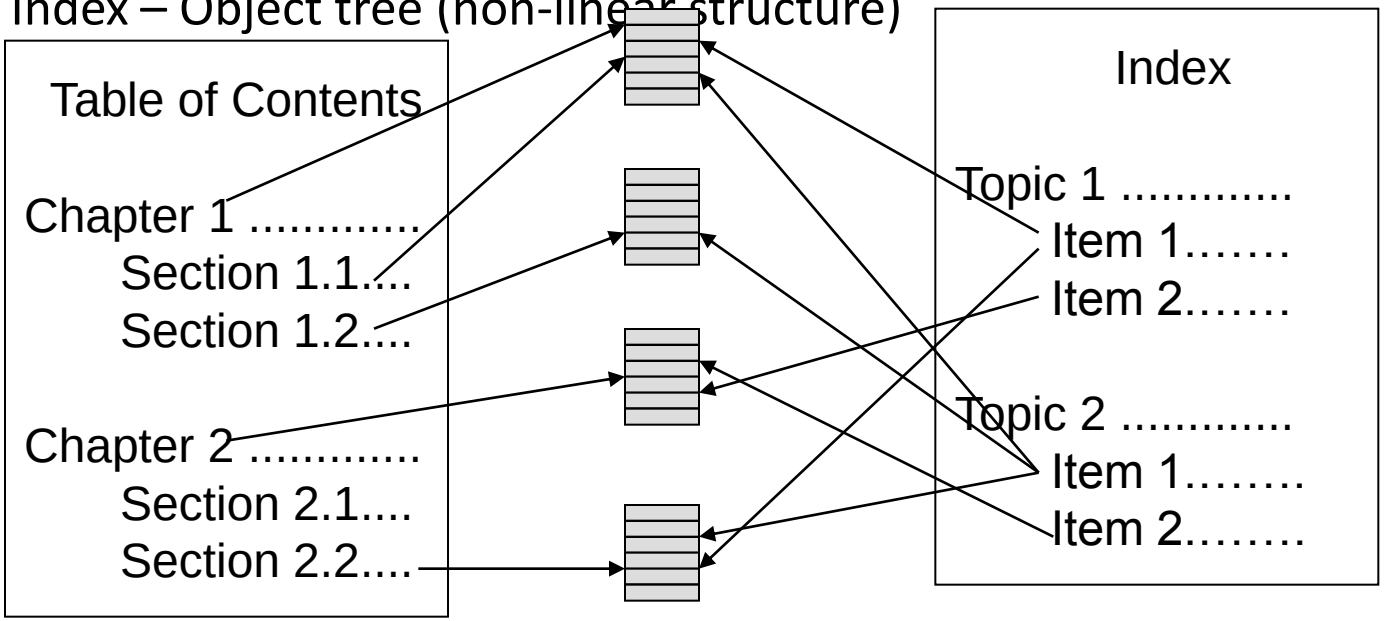


Content Management

- Creation & Classification
  - Title, textual annotation, creators, creation locations, and dates.
  - Categories such as genre, subject, purpose or language.
  - Review and guidance information: Age classification, parental guidance, and subjective review.
  - Related material information.
- Media coding, storage & file formats
  - Media profiles & master media
- Content Usage
  - Usage rights, usage record, and financial information

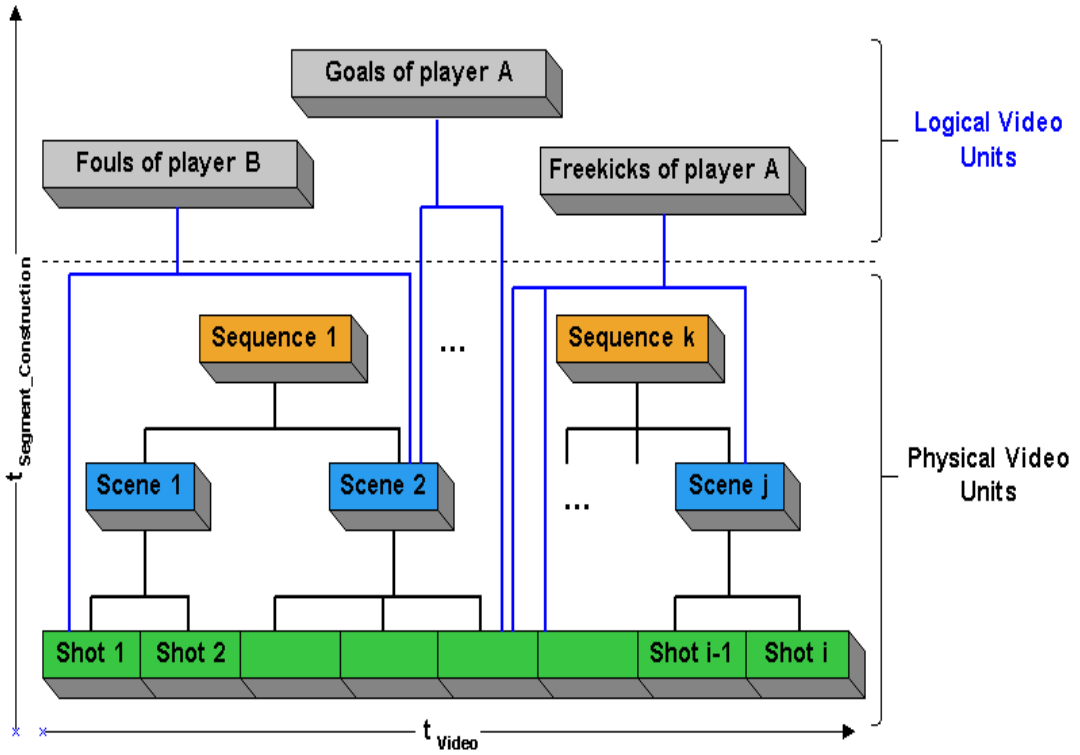
Content Description: Structural vs. Conceptual Aspects

- Program DS (in sense of TV program)
- Analogy to
  - Table of content – Region tree (linear partitioning)
  - Index – Object tree (non-linear structure)

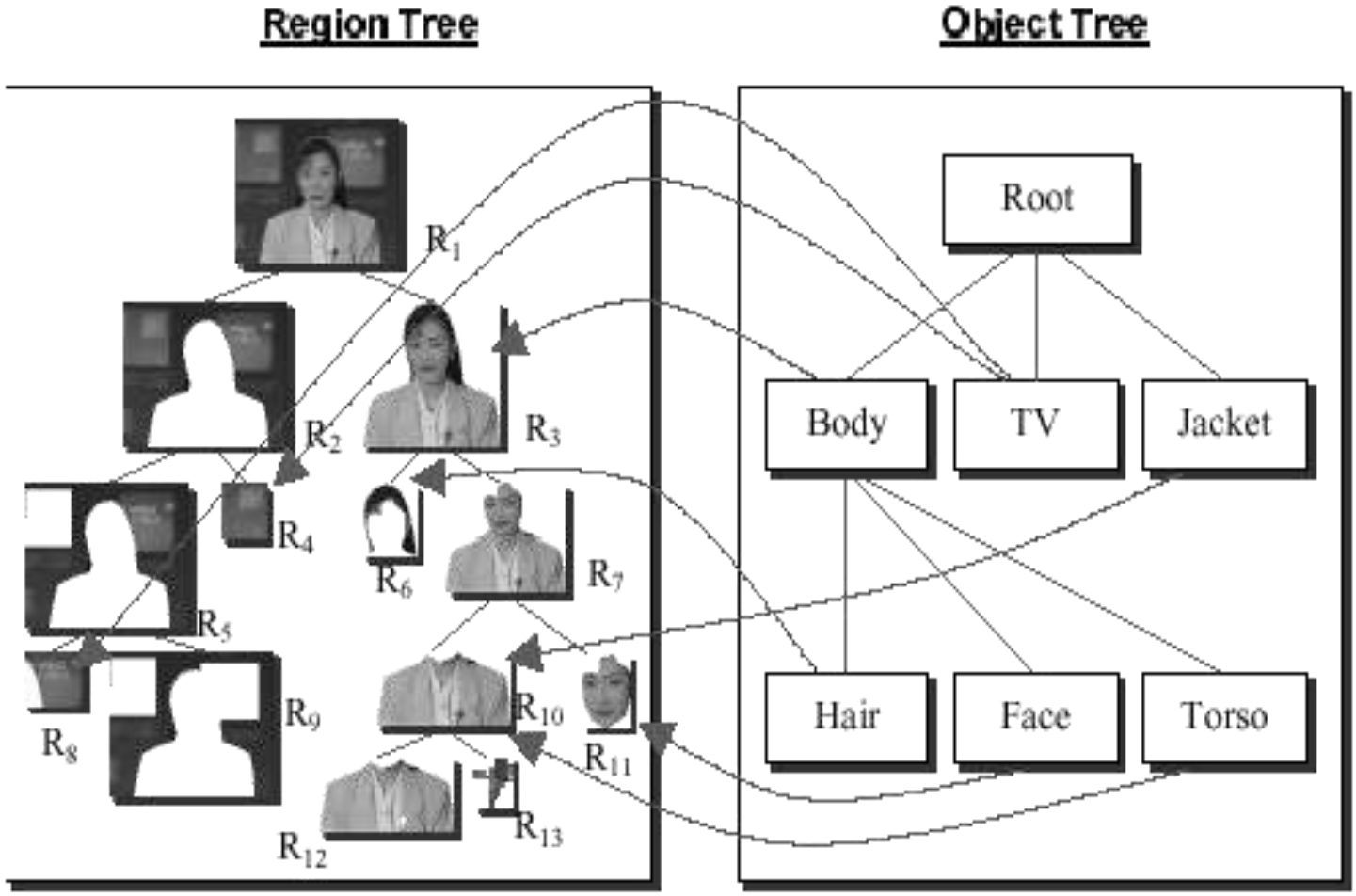


Content Description: Structural Aspects

- Divide a video stream into physical and logical video segments
- The higher the level of a physical video unit, the more semantic information is necessary
- Logical units are based on semantic content



Region and Object Trees



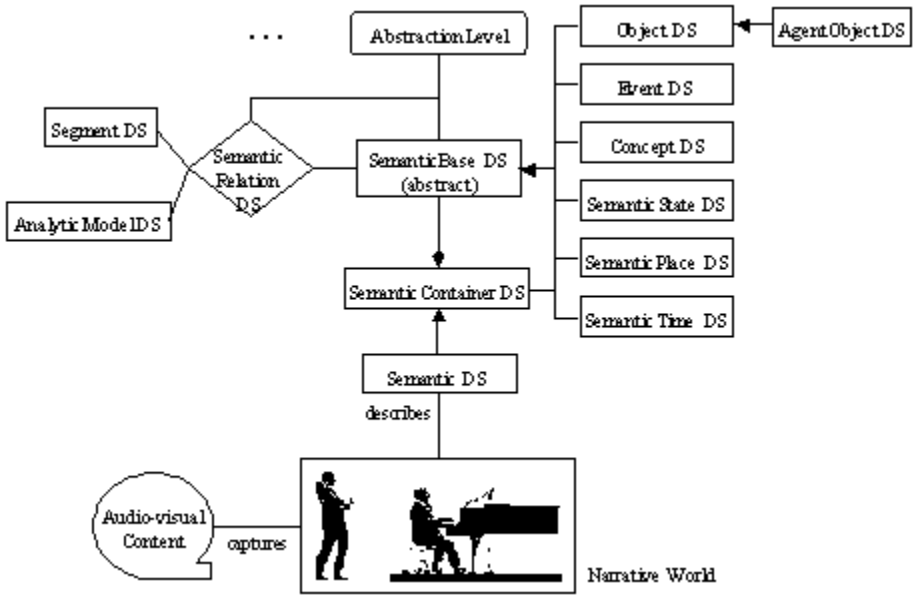
Content Description: Semantic Aspects

- Low Level Features
  - Extraction from Content
  - Descriptors for
    - Shape, color, texture (visual)
    - Timbre, rhythm (audio)
- High Level Features
  - Annotation
  - So called *semantic descriptors*
    - Textual information
    - Conceptual information

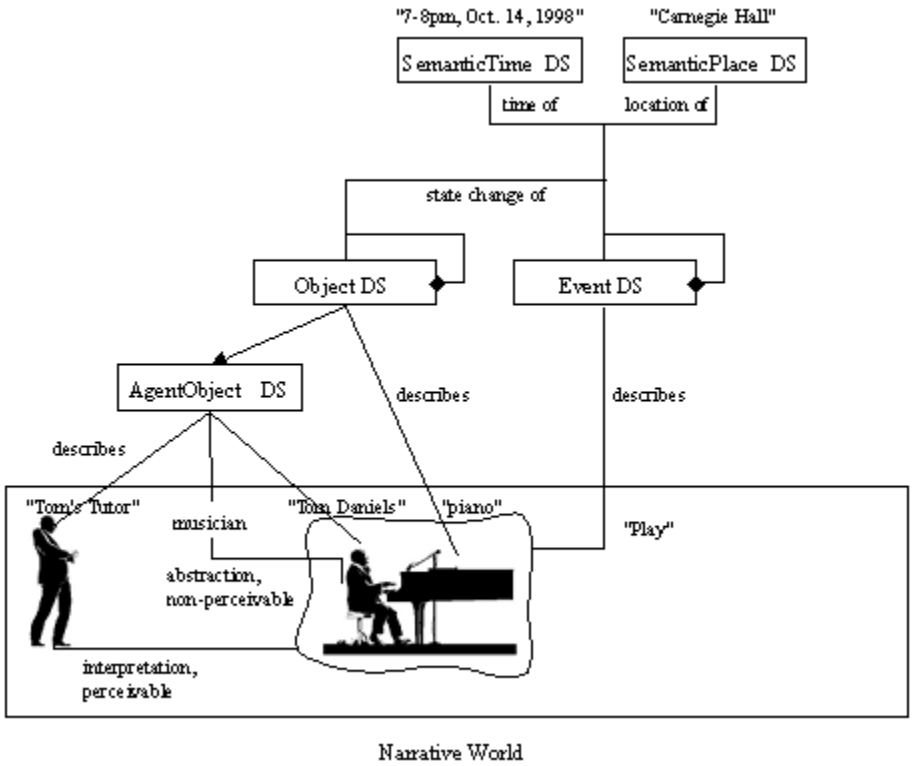
MPEG-7 High Level Descriptors

- Textual Descriptions
  - Text to describe temporal / spatial regions
- The W's
  - Structured way of textual descriptions
    - Who, Where, What Object, When, Why, How & Where
- Instead of textual descriptions
  - Controlled Terms
    - Dictionaries, Taxonomies, Classifications Schemes
  - Semantic Description Scheme

MPEG-7 Semantic Description Scheme



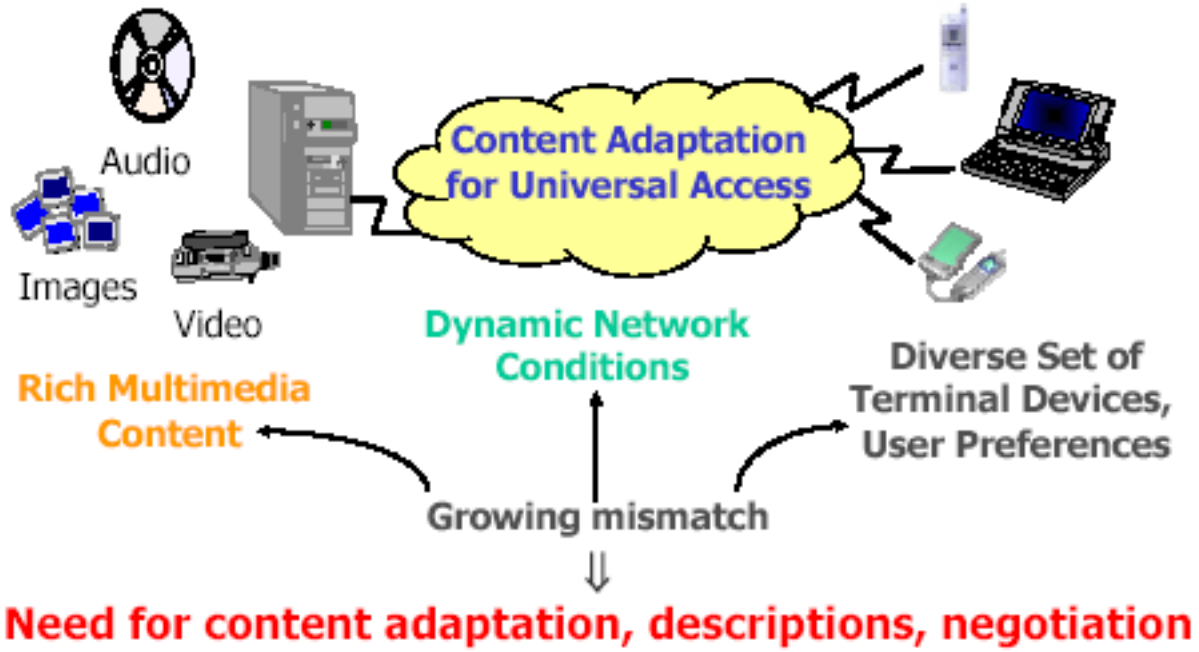
Actual Description in MPEG-7



Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- **MPEG-21**
- Metadata Generation & Annotation

MPEG-21 – motivation and scope



MPEG-21 Objectives

MPEG-21's goal is to create an *interoperable and integrated multimedia framework* in three steps:

1. **Develop "big picture"**: understand how the components of the framework are related and identify where gaps in the framework exist
2. **Fill the gaps**: develop new standard specifications where needed
3. **Integrate**: achieve the integration of standards to support harmonized technologies for the management of multimedia content

MPEG-21 Digital Item

A **Digital Item (DI)** is a structured digital object with a standard representation, identification, and metadata within the MPEG-21 framework

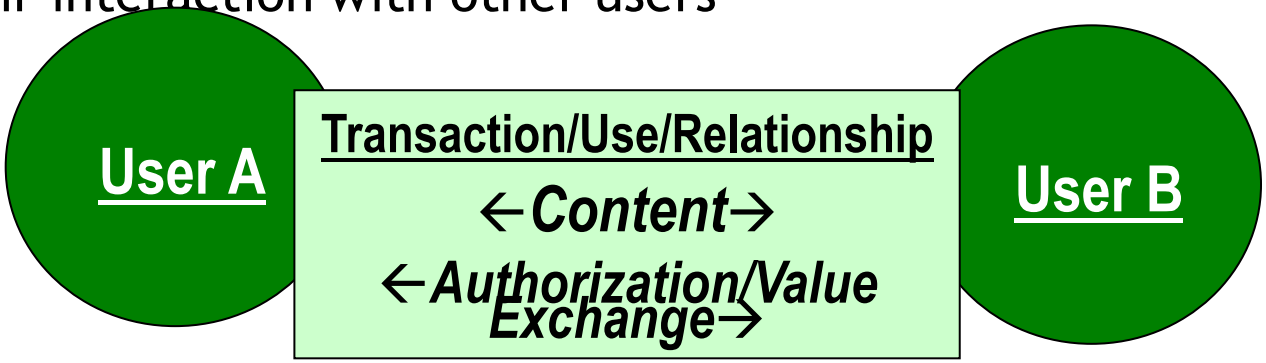
- Digital Items are “the content”
- DIs consist of
  - Resources (individual assets, distributed content),
  - Metadata (data about or pertaining the DI) and
  - Structure (relationships between parts of the DI)

Digital Item - Example

The DI is the fundamental unit for distribution and transaction within the MPEG-21 framework.

MPEG-21 User and User Interaction

- Any entity that interacts in the MPEG-21 environment or makes use of a Digital Item
- Users include individuals, organisations, corporations, consortia, governments, other standards bodies, etc.
- Roles including creators, consumers, rights holders, content providers, distributors, etc.
- Each User will assume specific rights and responsibilities according to their interaction with other users



Seven Architectural "Elements"

Digital Rights Management

- Pt. 4: **IPMP** Components
- Pt. 5: **R**ights **E**xpression **L**ang
- Pt. 6: **R**ights **D**ata **D**ictionary
- Amd.1: DII relationship types

Adaptation

- Pt. 7: **D**igital **I**tem **A**daptation
- Amd.1: Convers. And Permissions
- Amd.2: Dynamic and Distributed Adaptation

Processing

- Pt. 10: **D**igital **I**tem **P**rocessing
- Amd.1: Add'l C++ bindings

Systems

- Pt. 9: File Format
- Pt. 16: Binary Format
- Pt. 18: **D**igital **I**tem **S**treaming

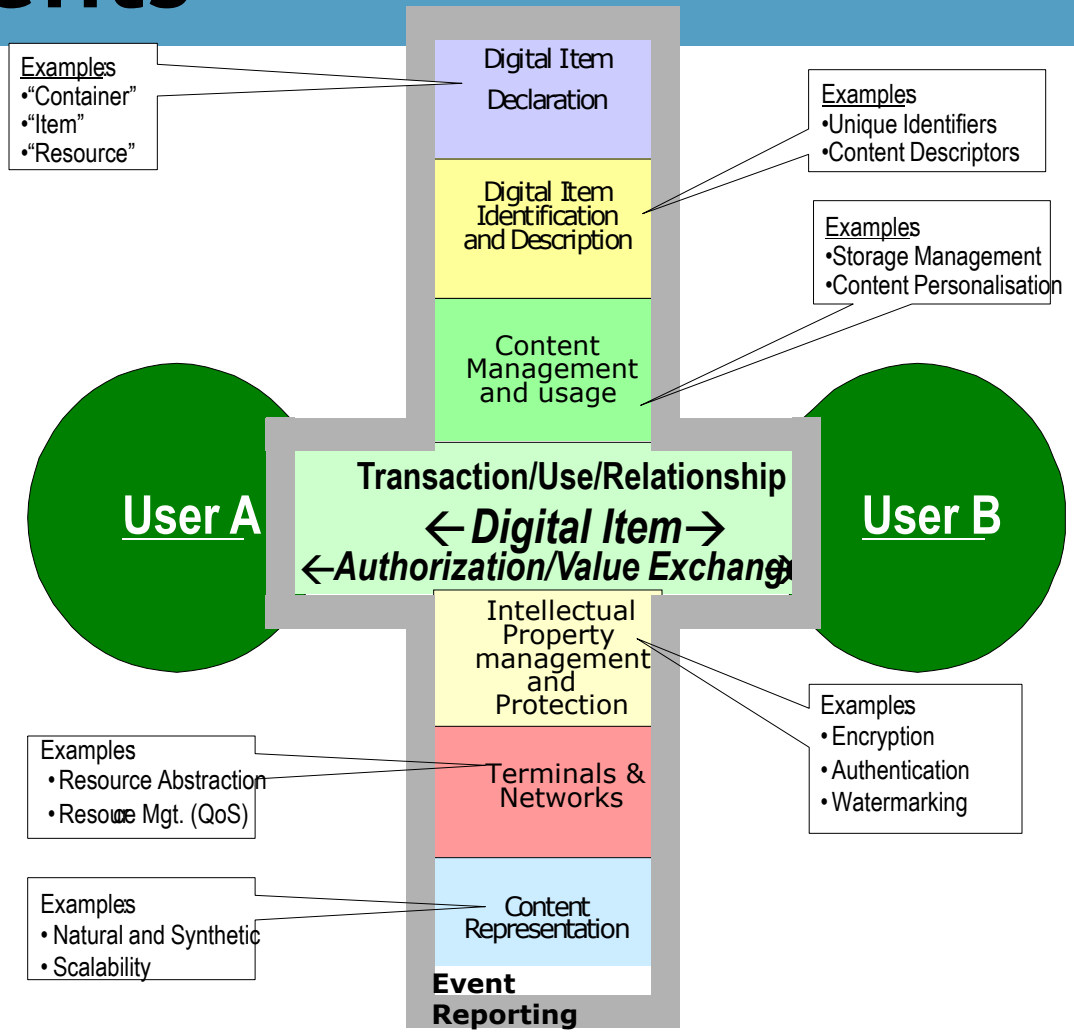
Misc

- Pt. 8: Reference Software
- Pt. 11: Persistent Association
- Pt. 12: Test Bed
- Pt. 14: Conform.
- Pt. 15: Event Reporting
- Pt. 17: Fragment Idenfication

Vision, Declaration, and Identification

- Pt. 1: Vision, Technologies and Strategy
- Pt. 2: **D**igital **I**tem **D**eclaration
- Pt. 3: **D**igital **I**tem **I**dentification

Roles of the Architectural Elements



Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- **Metadata Generation & Annotation**

Metadata Generation & Annotation

- Process of creating data about data
- Content has to be known
  - Watch & understand video / image collection
  - Listen and assess audio
- Metadata standard has to be known
  - What are the possible fields
  - What are the used classification systems

Evaluation (1/2)

- Goal: Identify the opinion of users on manual semantic annotation

- 5 Users with following (median) background:
  - 17 years of computer experience
  - Using a computer 50 h / week
  - 2 years experience with digital photo cameras
  - 4 years experience with imaging software

Evaluation (2/2)

- 2 Tasks were given:
  - Annotate a photo with a given description and an extensive prior introduction to semantic photo annotation with Caliph
    - video was shown,
    - concept was explained and
    - questions were answered
  - Annotate a photo fully on your own
  - After Tasks:
    - Questionnaire with several subjective questions
    - Evaluation Scale from: -3 (strongly disagree) to 3 (strongly agree)

Evaluation Results: General Questions

- The concept of meta data is very new to me: **-2.6**
- It was easy to understand the concept of semantic meta data while using Caliph: **1.8**
- The visualization of the semantic meta data within Caliph is easy to understand and interpret: **2.2**
- The annotation of images with textual descriptions can be done fast and easily: **1.4**
- The annotation of images with semantic meta data can be done fast and easily: **1.2**
- I can see an obvious benefit by using semantic meta data for image (multimedia) annotation: **1.4**

Scale: (disagree) -3 to 3 (agree)

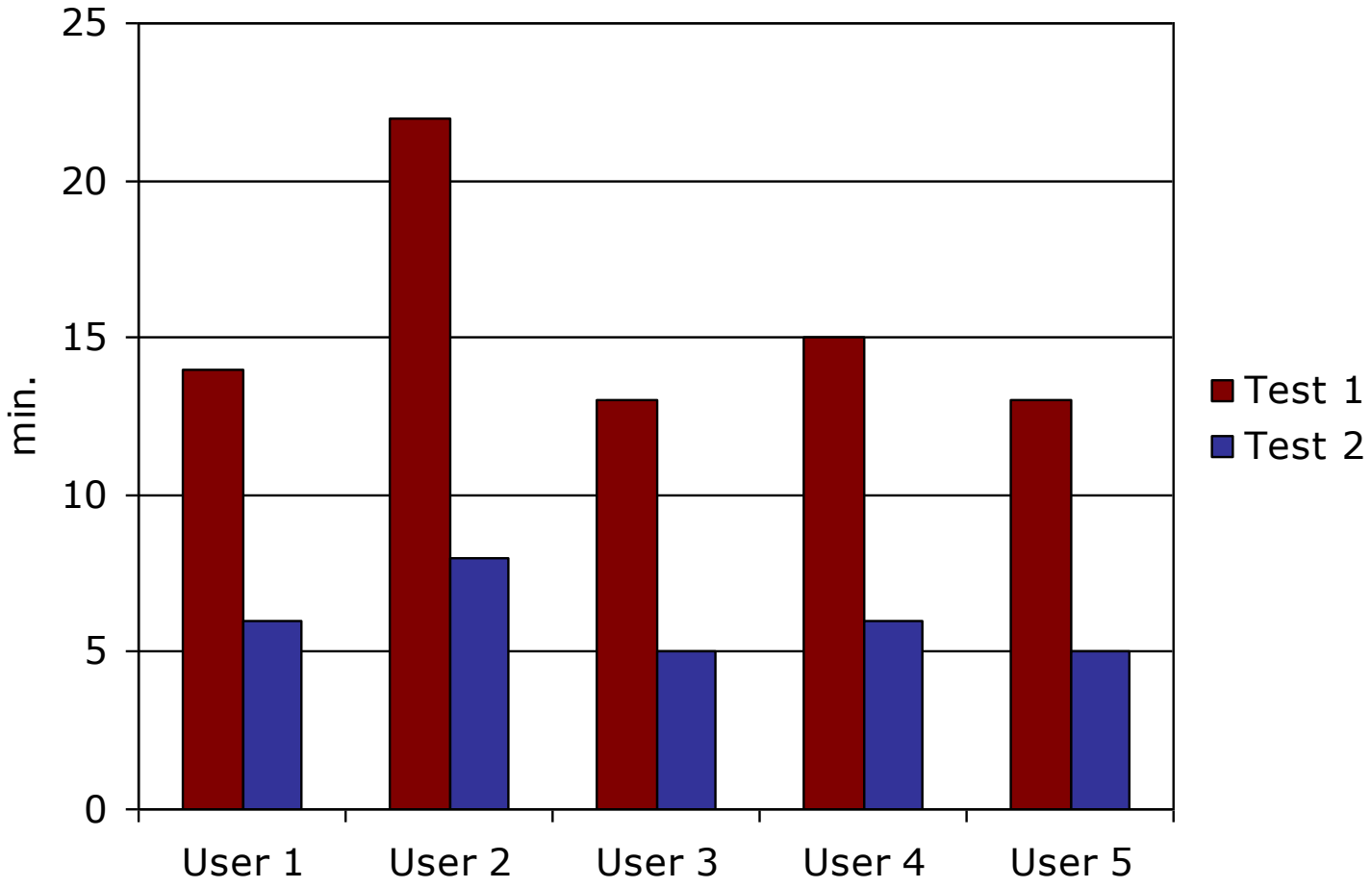
Evaluation Results: Scenario based questions

1. The complexity of semantic annotation is too high to be useful for organizing photos.
2. I would find it easy to annotate a large set digital photos (e.g. 100+).
3. I would recommend Caliph or a similar tool to annotate digital photos.
4. I can see an obvious benefit by using semantic meta data for the organization of photos.

Personal	Newspaper
-0.6	-1.8
-0.6	-0.2
0.8	1.4
1.4	2.2

Scale: (disagree) -3 to 3 (agree)

Evaluation Results: Annotation performance



Evaluation Results: Annotation performance

- Median times for annotation:
  - 15.4 minutes for the 1st test and
  - 6 minutes for the 2nd test
- Median time in a self test with 17 photos:
  - 1 minute and 53 seconds per photo
- That results in an approximate time of 10 h 27 min. for annotation of a set of 333 photos

Evaluation Results: Diversity of Annotations (2nd test)

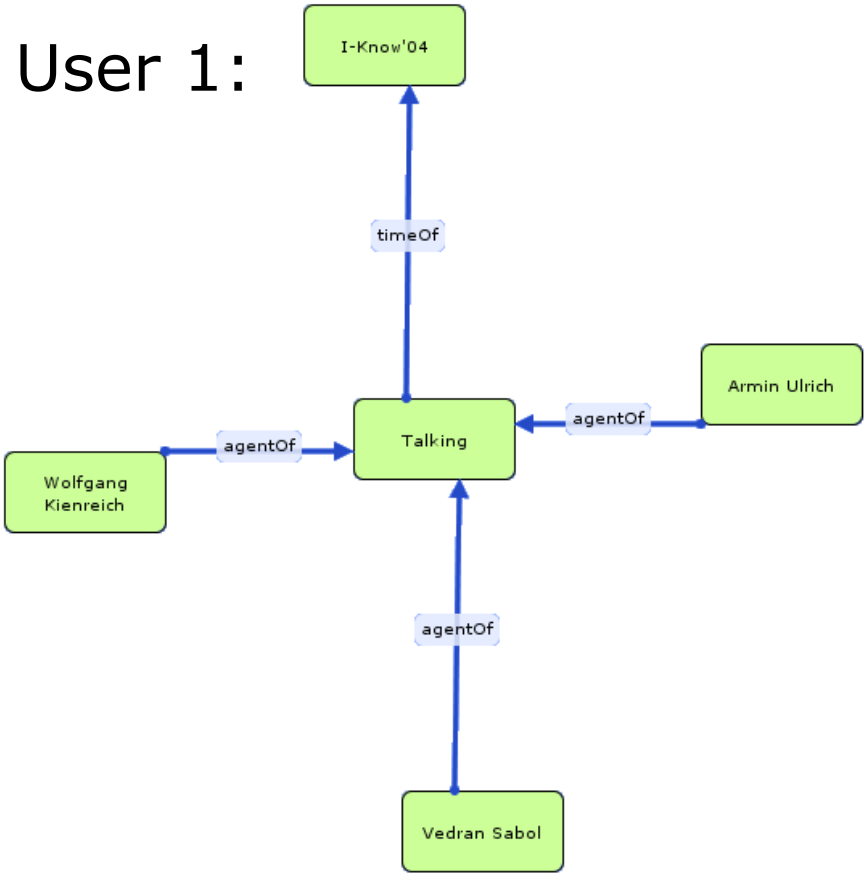
- Structured text annotation field “Who”:
  1. Vedran, Wolfgang, Armin
  2. Wolf, Armin, Vedran
  3. Wolfgang Kienreich, Vedran Sabol, Armin Ulbrich
  4. wolfgang, armin, vedran
  5. W.Kienreich,A.Ulbrich,V.Sabol

Evaluation Results: Diversity of Annotations (2nd test)

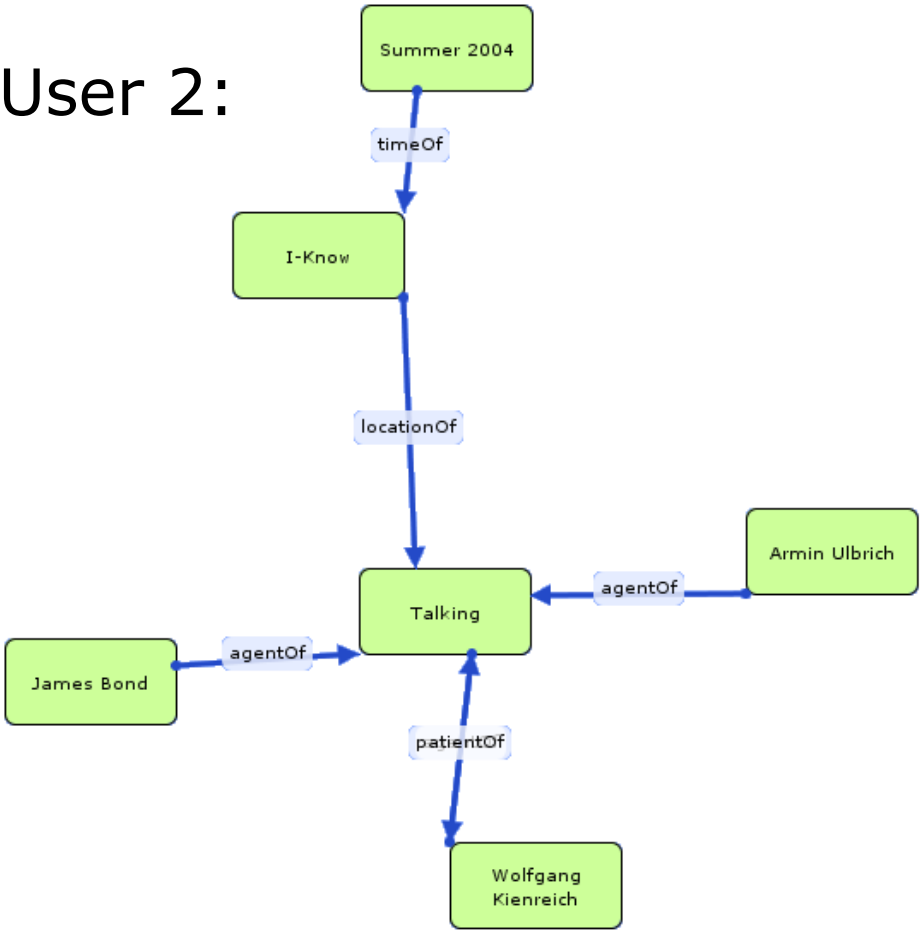
- Free text annotation:
  1. Stadthalle, Graz, Austria I-Know '04 Knowledge Managment Conference
  2. The three are sitting ...
  3. Wolfgang Kienreich, Armin Ulbrich und Vedran Sabol (v.l.n.r.) sprechen miteinander auf der I-Know.Wolfgang Kienreich, Vedran Sabol, Armin Ulbrich are at I-Know, Graz for Talking
  4. Stadthalle, Graz, Austria I-Know '04 Knowledge Managment Conference
  5. Wolfgang,Armin and Vedran talking to each other on I-Know 04 at Stadthalle Graz.

Evaluation Results: Diversity of Annotations (2nd test)

User 1:

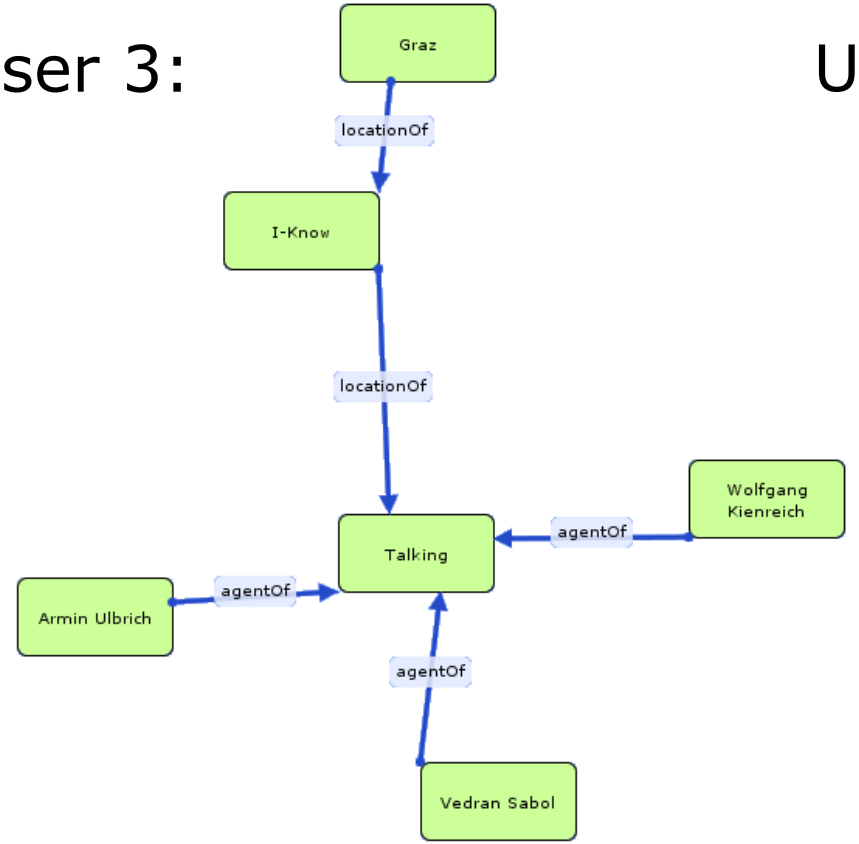


User 2:

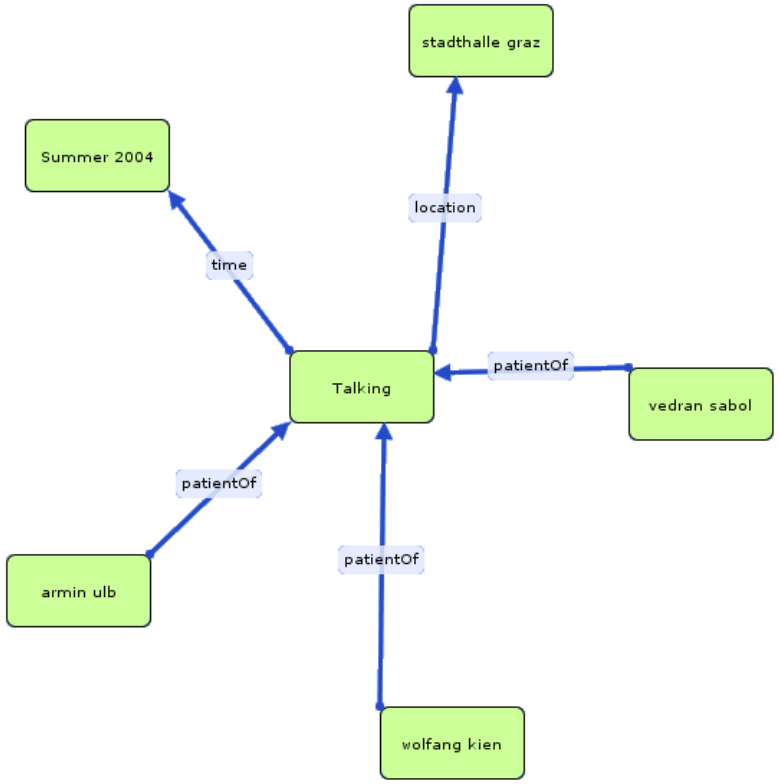


Evaluation Results: Diversity of Annotations (2nd test)

User 3:

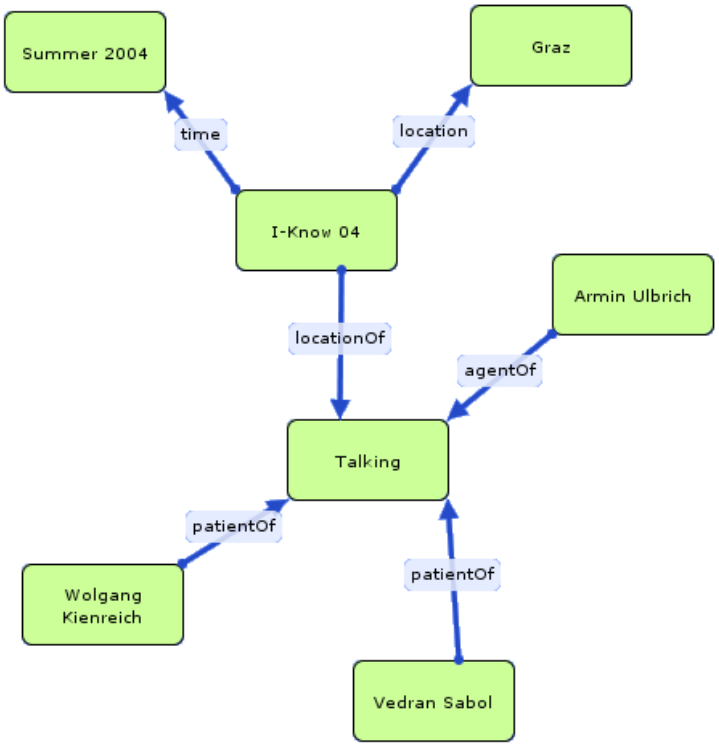


User 4:



Evaluation Results: Diversity of Annotations (2nd test)

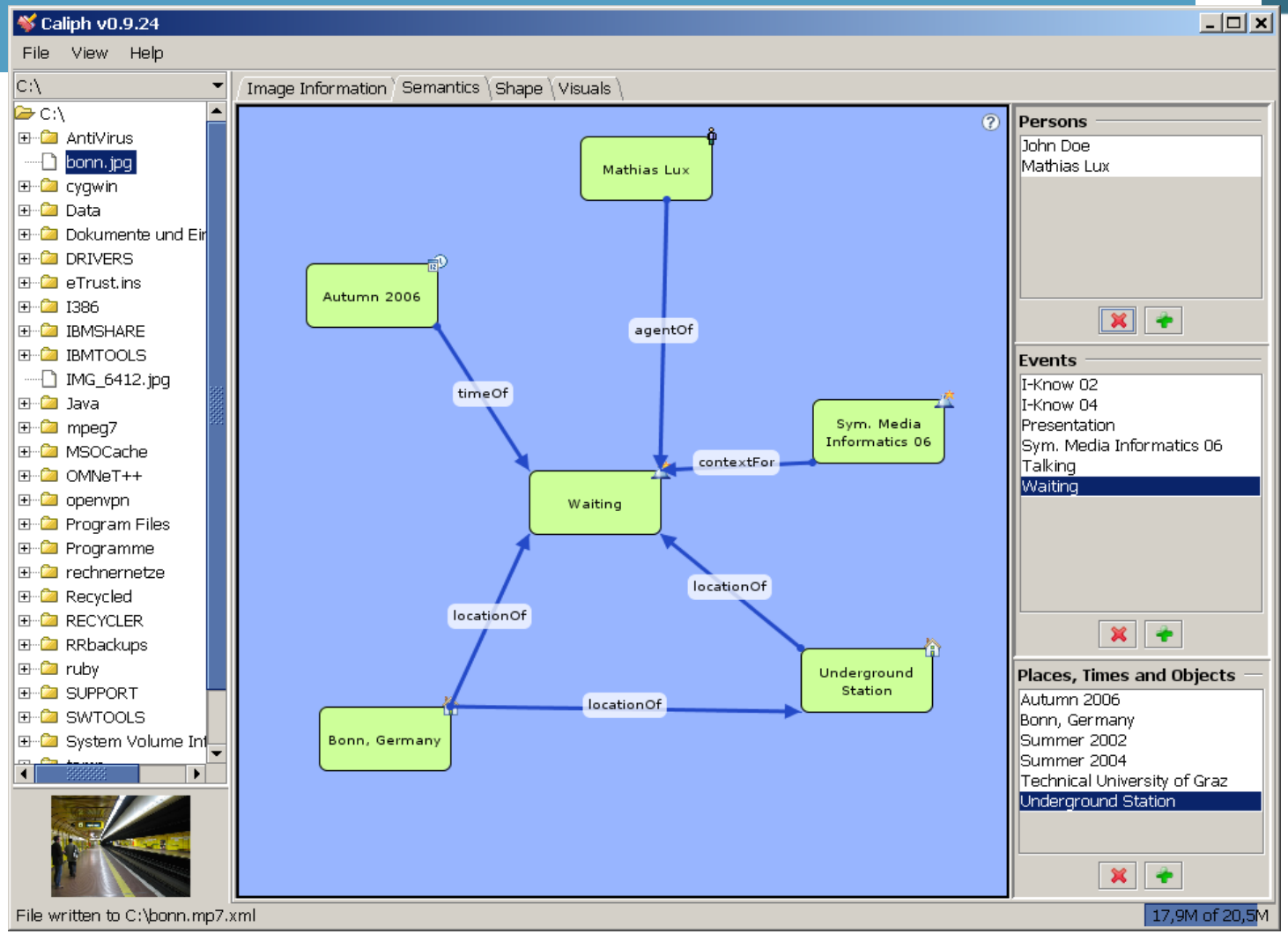
User 5:



Lessons Learned

- Users like the graphical annotations editor
- Users see semantic annotation in a professional (business) environment
- Semantic annotation is very time consuming
- The MPEG-7 nomenclature is not intuitive
  - Semantic agent / place / object & relations
  - Creator of image / description / quality rating
- Tagging with central tag repository …

Demo


Caliph v0.9.24
File View Help
C:\
Image Information Semantics Shape Visuals
AntiVirus
bonn.jpg
cygwin
Data
DRIVERS
eTrust.ins
I386
IBMSHARE
IBMTOOLS
IMG_6412.jpg
Java
mpeg7
MSOCache
OMNeT++
openvpn
Program Files
Programme
rechnernetze
Recycled
RECYCLER
RRbackups
ruby
SUPPORT
SWTOOLS
Mathias Lux
Autumn 2006
agentOf
timeOf
Sym. Media Informatics 06
contextFor
Waiting
locationOf
locationOf
Underground Station
locationOf
Bonn, Germany
Persons
John Doe
Mathias Lux
Events
I-Know 02
I-Know 04
Presentation
Sym. Media Informatics 06
Talking
Waiting
Places, Times and Objects
Autumn 2006
Bonn, Germany
Summer 2002
Summer 2004
Technical University of Graz
Underground Station
File written to C:\bonn.mp7.xml
17,9M of 20,5M