

# Multimedia and Semantic Metadata



*Grundlagen des Wissensmanagement  
WS 2007/2008*

Mathias Lux, [mlux@itec.uni-klu.ac.at](mailto:mlux@itec.uni-klu.ac.at)



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 2.0 License. See <http://creativecommons.org/licenses/by-nc-sa/2.0/at/>



Department for Information Technology, Klagenfurt University, Austria

## Mathias Lux



<http://www.uni-klu.ac.at>

- Techn. Mathematik / TU Graz
- Telematik / TU Graz
- Know-Center & KMI
- Alpen Adria Universität Klagenfurt
  - Information Technology – ITEC
  - Multimedia (mainly video), Networks, ...



ITEC, Klagenfurt University, Austria

2

# Agenda



<http://www.uni-klu.ac.at>

- What is Multimedia?
- Multimedia Metadata
  - Low vs. High level features
  - Standards
  - Annotation Issues
- Some Conclusions

# What is Multimedia?



<http://www.uni-klu.ac.at>

- Combination of different media
  - Audio + Video
  - Audio + Video + Text
- Research areas
  - Coding & Compression
  - Transmission & Delivery
  - Retrieval & Management
  - Perception, Interfaces & Interactivity
  - etc.

## Multimedia Metadata



<http://www.uni-klu.ac.at>

- Motivation & general aspects
- Low level vs. high level features
- Common Formats
  - Media Production
  - Ontologies
  - Home User
  - MPEG-7
- Multimedia Annotation

## Challenges in Multimedia



<http://www.uni-klu.ac.at>

- Consumer needs help in
  - Finding the proper data
  - Filtering the not required data
- Consumption constraints
  - Device capabilities
  - Digital rights & billing
- Processing online or in real time
  - Search & Filtering: Too slow
  - Transmission & Adaptation: Limited

## Multimedia Metadata:



<http://www.uni-klu.ac.at>

- Consumer needs help in
  - Finding & Filtering: **Indexing & Classification**
- Consumption constraints
  - Device capabilities: **Interoperable Descriptions**
  - Rights & billing: **Interoperable Processes & Formats**
- Processing online or in real time
  - Search & Filtering: **Based on Metadata & Text**
  - Adaptation: **Distributed, Based on Metadata**

## Metadata Problems



<http://www.uni-klu.ac.at>

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

# Aspects of Metadata



<http://www.uni-klu.ac.at>

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

# Aspects of Metadata: Content Description



<http://www.uni-klu.ac.at>

- Agenda
  - Overview about a presentation or a sequence of information to a particular topic
- Table of Contents
  - A list of all segments and their position
- Abstract
  - Describes the topic of a content within a few sentences.
- Preface
  - Some words of the author
- Structure
  - For consumption & navigation

## Aspects of Metadata: Content Description



<http://www.uni-klu.ac.at>

- Keywords & index
  - Content description and lookup of concepts
- Summary
  - Overview of the most important aspects of an content or its deductions, e.g. video summary
- References & footnotes
  - Additional material, sequels, ...
- Comments
  - For interactive environments
- Categories
  - Conceptual classification in taxonomies (genre, audience etc.)
- Languages
  - Which languages are used / available

## Aspects of Metadata: Administrative Metadata



<http://www.uni-klu.ac.at>

- Associated persons
  - Authors: Content creators
  - Contributors: People who contributed to the content
- History of Changes
  - Changes in content and metadata with author, date, position in the content and sort of action
  - Especially in production
- Unique identifier
  - e.g. URI or database id
- Versions
  - Versioning information related to the history

## Aspects of Metadata: Quality Aspects



- Weight
  - Prioritization of segments, e.g. scenes
- Expiration Date
  - Time period of validity of the content
- Recessions
  - Opinions, arguments from others
- Process description & history
  - Who corrected, translated and approved the content e.g. within an workflow
- Quality Assessment
  - Rating of the (e.g. visual) quality of the content

## Aspects of Metadata: Legal Metadata



- Copyright
  - Person or company legally permitted to sell or trade with the content.
- Publishing Date
  - Date when the content has been released to public.
- License Model
  - This is the mode how consumers are allowed to reuse the content

## Aspects of Metadata: Technical Metadata



<http://www.uni-klu.ac.at>

- Standards
  - Description of the standard for storage / transmission
- Application/System
  - Tools for content and metadata processing
- Resolution & compression
  - Note: Compression & container are different aspects
- Encryption Method
  - In case of encrypted content / DRM
- Storage Media
  - CDs, tapes, MO, paper, HDD, etc.
- Logs
  - Technical history

## Multimedia Metadata



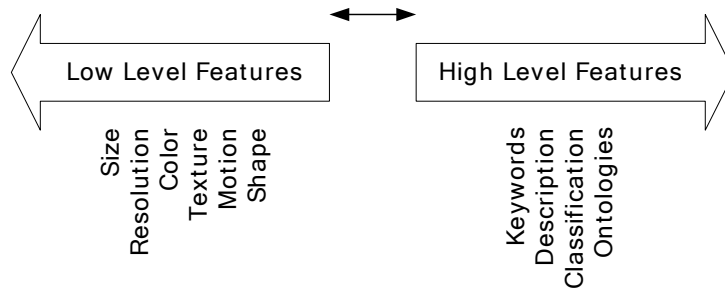
<http://www.uni-klu.ac.at>

- Motivation & general aspects
- **Low level vs. high level features**
- Common Formats
  - Media Production
  - Ontologies
  - Home User
  - MPEG-7
- Multimedia Annotation

# Multimedia Metadata



<http://www.uni-klu.ac.at>



Semanti

# Low Level Features



<http://www.uni-klu.ac.at>

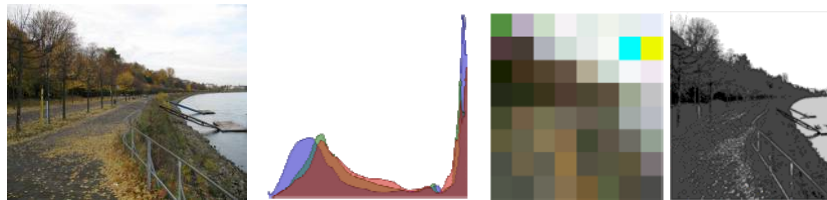
- Automated creation (extraction)
- Global vs. local features
- Common Types:
  - Color (Color histograms, dominant colors)
  - Texture (Frequency layouts, edge histograms)
  - Motion (Motion trajectories, motion activities)
  - Shape

# Color Features



<http://www.uni-klu.ac.at>

- Color Histogram
  - Quantity of Colors (Ranges) compared
  - Transformation to a “better” Color Space
- Dominant Color
  - Only the most prominent Colors are used
- Color Distribution
  - Which colors appear where in the image?



# Texture



<http://www.uni-klu.ac.at>

- Edge Detection & Description
  - Edges in sub-parts of images
  - left to right, top to bottom, diagonal (2x), non directional
- Overall Texture Properties
  - Coarseness
  - Regularity



## Motion & Shape



<http://www.uni-klu.ac.at>

- Camera Motion
  - Tilt, Pan, Rotate, Zoom
- Object (Background) Motion
- Shape descriptions
  - Connectivity, Number of Holes
  - Area, Perimeter, Diameter
  - Eventually based on the convex hull



## High-level features



<http://www.uni-klu.ac.at>

- Created (mostly) manually
  - Segmentation, Description, Tagging, ...
  - Time-consuming and expensive task
- Distinction physical vs. logical segments
- Semantic ambiguity
  - biased subjective interpretations of content
- Based on a *narrative world*
  - Context for interpretation

## Semantics & High Level Metadata



<http://www.uni-klu.ac.at>

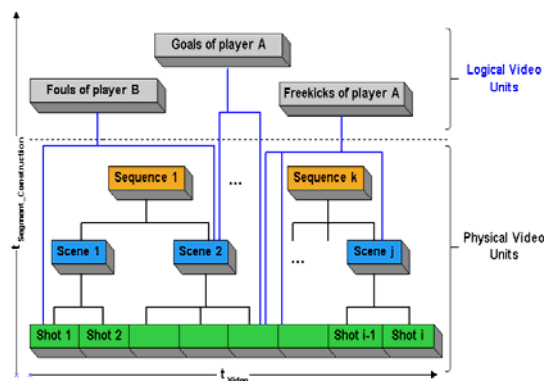
- Interpretation of “semantics” depends on context
- Different points of view
  - All high level features are semantics
  - Features that cannot be extracted automatically are semantic features
  - All associated text describing the content
  - Features capturing the semantics of the content
  - Clearly defined concepts and their relations to the content (taxonomies, ontologies, ...)

## Example: Video Segmentation



<http://www.uni-klu.ac.at>

- 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



## Example: Video Segmentation



<http://www.uni-klu.ac.at>

- Physical units:
  - Scenes are not merely technical issues
  - Shots are defined but not non-ambiguous
- Logical units
  - Clear use of concepts
  - But no obvious relations
    - Freekick of Player C is result of Foul of Player B ...

## Multimedia Metadata



<http://www.uni-klu.ac.at>

- Motivation & general aspects
- Low level vs. high level features
- **Common Formats**
  - Media Production
  - Ontologies
  - Home User
  - MPEG-7
- Multimedia Annotation

## Media Production: Dublin Core



<http://www.uni-klu.ac.at>

- 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



<http://www.uni-klu.ac.at>

- Aims to provide ...
  - a universal standard for metadata exchanges between professional media organizations
  - a definition of common meaning to the data fields and values that most broadcasters use in order to enable exchanges
  - 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



<http://www.uni-klu.ac.at>

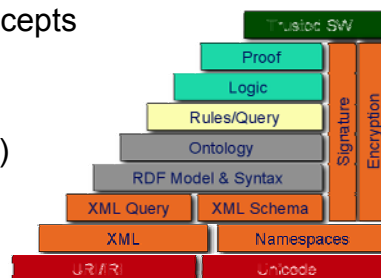
- SMPTE Metadata Dictionary
  - Society of Motion Picture and Television Engineers
    - Since 1916, 61 members
  - Standard for metadata exchange in TV
  - Defines set of attributes / fields
- MXF DMS-1
  - Metadata bundled with the Material Exchange Format (MXF)
  - Open format for the broadcasting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

## Ontologies: RDF



<http://www.uni-klu.ac.at>

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



## Ontologies: SKOS



<http://www.uni-klu.ac.at>

- 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



<http://www.uni-klu.ac.at>

- Multimedia Semantics : Incubator Activity of the W3C

### **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.

## Home User



<http://www.uni-klu.ac.at>

- Exchangeable Image File Format (EXIF)
  - Japan Electronic and Information Technology Industries Association (JEITA)
  - Extensive format for technical aspects
  - Settings and sensor reading at the time of recording
  - Mostly images from digital cameras
- IPTC Information Interchange Model (IIM)
  - Several elements to describe images (assets)
  - Spread by the use within Adobe applications

## Home User



<http://www.uni-klu.ac.at>

- 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



<http://www.uni-klu.ac.at>

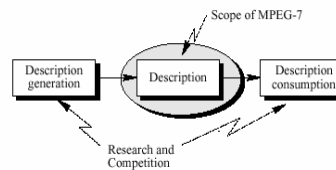
- Electronic Program Guide (EPG)
  - In use in conjunction with DVB
  - Simple format in additional stream
- Multimedia Home Platform (MHP)
  - In use in Austrian DVB-T
  - Proprietary format for data + function
  - Based on Java

## The MPEG-7 standard



<http://www.uni-klu.ac.at>

- Make searching for multimedia content as easy as searching for text is today
- Interoperable management of A/V data, such as
  - Searching
  - Filtering
  - Indexing
  - Accessing
- Associates descriptions (meta data) with content
  - Format of the descriptions must be standardized
  - Generation and consumption of those generally not

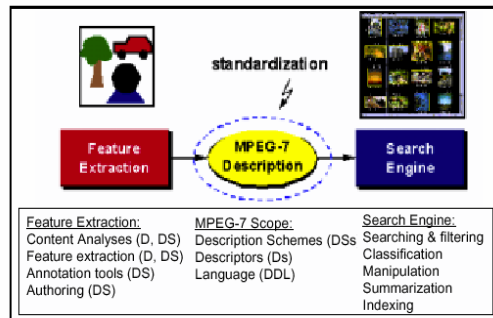


# Kinds of descriptions



http://www.uni-klu.ac.at

- Information about the content
  - Title, author, recording date, copyright, coding format ...
- Information extracted from the content
  - Combination of low and high level descriptors
- Forms of descriptions
  - Textual (XML document)
  - Binary Format for MPEG-7 (BiM)



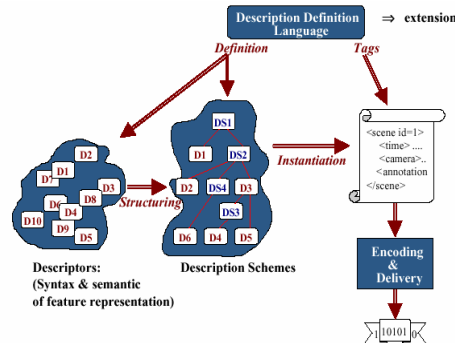
uni@klu

# Elements of the MPEG-7 standard



http://www.uni-klu.ac.at

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



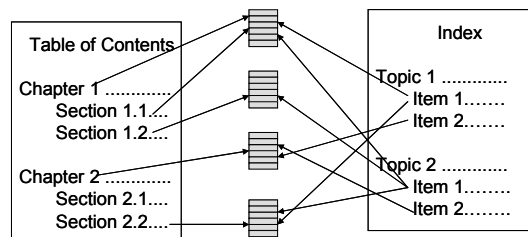
uni@klu ITEC, Klagenfurt University, Austria

# Structural vs. conceptual aspects



<http://www.uni-klu.ac.at>

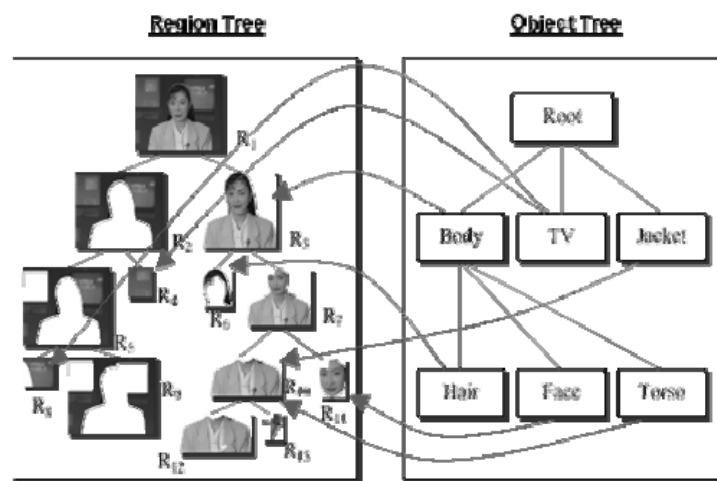
- Program DS (in sense of TV program)
- Analogy to
  - Table of content – Region tree (linear partitioning)
  - Index – Object tree (non-linear structure)



# Region and object trees



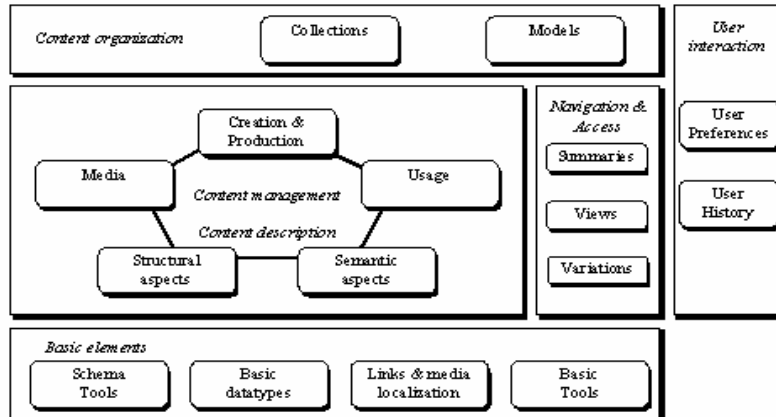
<http://www.uni-klu.ac.at>



# Scope of MPEG-7



<http://www.uni-klu.ac.at>



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

# MPEG-7 High Level Descriptors



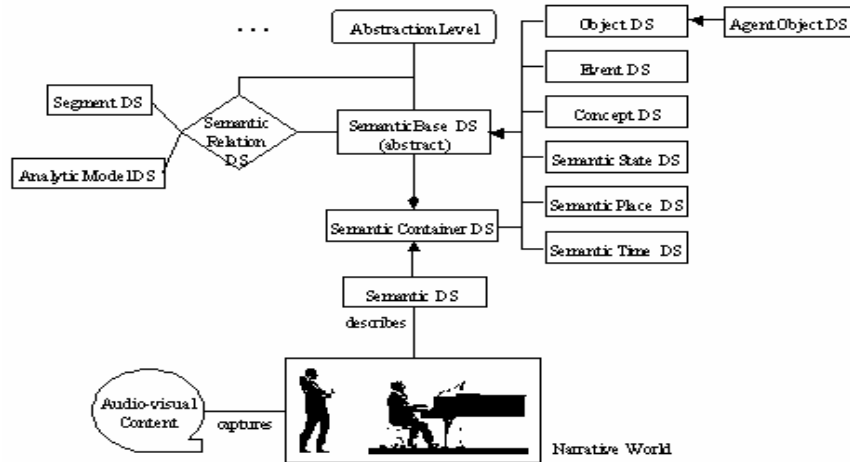
<http://www.uni-klu.ac.at>

- Textual Descriptions
  - text to describe temporal / spatial regions
- The W's
  - Structure 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



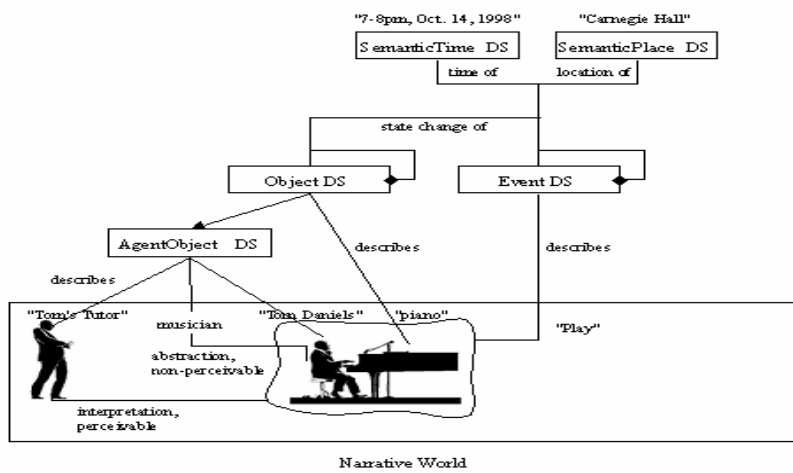
<http://www.uni-klu.ac.at>



# Actual Description in MPEG-7



<http://www.uni-klu.ac.at>



## RDF vs. MPEG-7



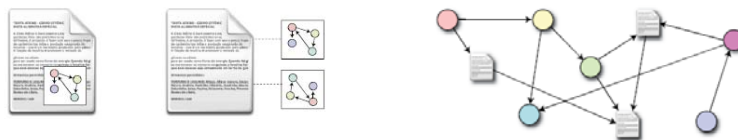
<http://www.uni-klu.ac.at>

- Distinguish between structure (triples) and content (values, e.g. agent description)
- Distributed approach
  - Triples are “stand alone” information
  - Although XML cannot provide this functionality
  - MPEG-7 is ‘one document’
- MPEG-7 documents can be transformed to RDF (Hunter 2001)

## Document vs. Metadata centric approach



<http://www.uni-klu.ac.at>



## Conceptual Graphs vs. MPEG-7



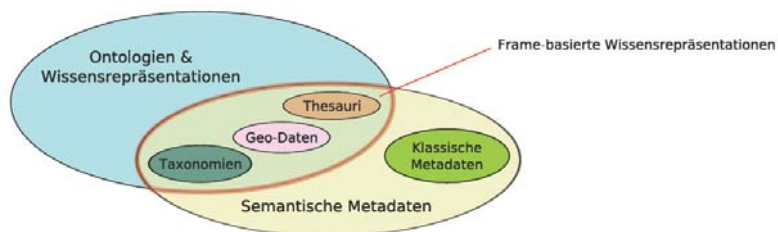
<http://www.uni-klu.ac.at>

- Introduced by J. Sowa
  - Mainly to describe data bases
- Conceptual Graphs describe a state
  - E.g. “Green grass grows slowly”
  - Can be “ill-formed”
- Conceptual Graphs can be used to describe the “state of multimedia content”
- RDF documents can be transformed to a Conceptual Graph (Corby, Dieng & Hebert 2000)

## Ontologies & Semantic Metadata



<http://www.uni-klu.ac.at>



## Summary: Semantic Metadata & Standards



- KOS are in use for archiving multimedia data (DC, KOS, ID3, ...)
- W3C investigates possibilities to annotate images with OWL
- MPEG has defined semantic descriptors in MPEG-7
- RDF, Conceptual Graphs and MPEG-7 are strongly related

## Multimedia Metadata



- Motivation & general aspects
- Low level vs. high level features
- Common Formats
  - Media Production
  - Ontologies
  - Home User
  - MPEG-7
- **Multimedia 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 / day
  - 2 years experience with digital photo cameras
  - 4 years experience with imaging software

## Evaluation (2/2)



<http://www.uni-klu.ac.at>

- 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
  - Evaluation Scale from: -3 (strongly disagree) to 3 (strongly agree)

## Evaluation Results: General Questions



<http://www.uni-klu.ac.at>

- 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



<http://www.uni-klu.ac.at>

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.

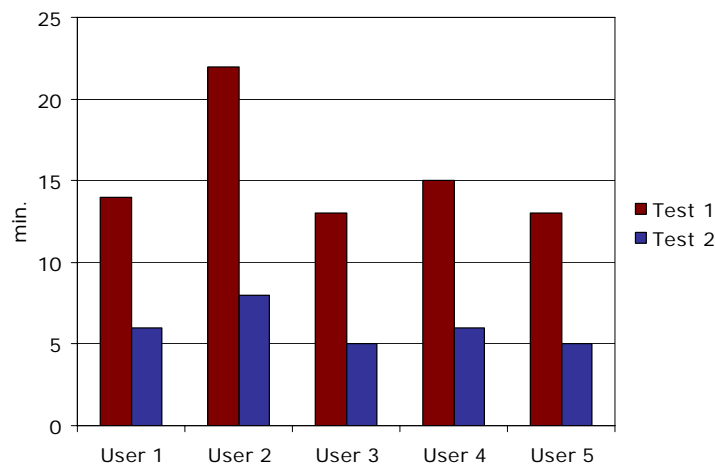
Personal	Newspaper
-0.6	-1.8
-0.6	-0.2
0.8	1.4
1.4	2.2

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

## Evaluation Results: Annotation performance



<http://www.uni-klu.ac.at>



## Evaluation Results: Annotation performance



- Median times for annotation:
  - 15.4 minutes for the 1<sup>st</sup> test and
  - 6 minutes for the 2<sup>nd</sup> 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 (2<sup>nd</sup> 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 (2<sup>nd</sup> test)



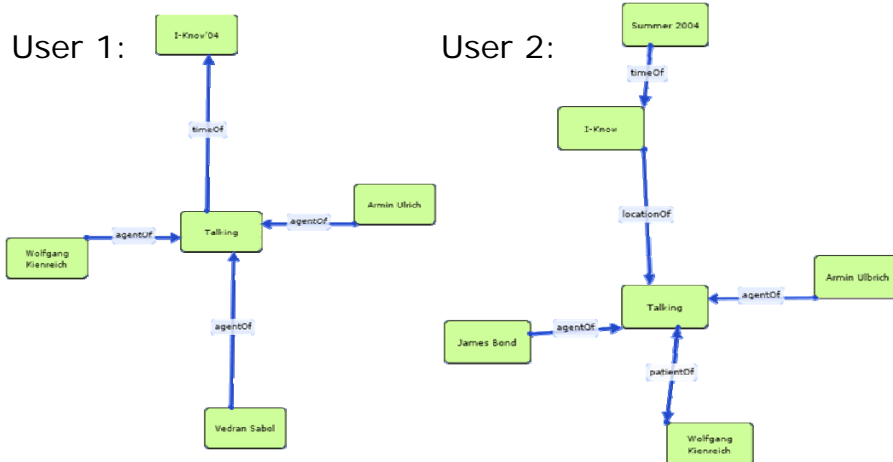
<http://www.uni-klu.ac.at>

- Free text annotation:
  1. Stadthalle, Graz, Austria I-Know '04 Knowledge Management 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 Management Conference
  5. Wolfgang, Armin and Vedran talking to each other on I-Know 04 at Stadthalle Graz.

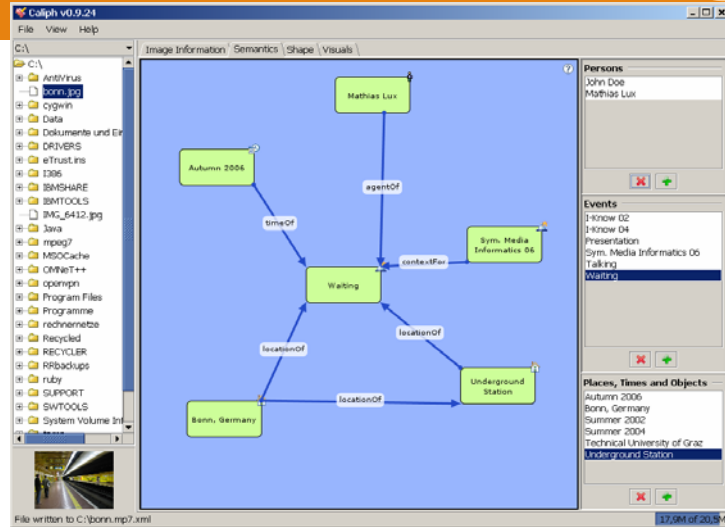
## Evaluation Results: Diversity of Annotations (2<sup>nd</sup> test)



<http://www.uni-klu.ac.at>



# Demo



# Conclusions

- Available metadata standards are diverse
  - Simple in means of usability & features
  - Broad, complex and powerful
- The concept of “semantics” is interpreted in different ways.
- Semantics (Meaning) is hidden even in low level features.
- Practical use of fine grained descriptions in multimedia is sparse.