



Caliph & Emir



Semantics in Multimedia Retrieval and Annotation

Mathias Lux
mlux@know-center.at

<http://www.know-center.at>

Overview

- Introduction
- Related Work
- Annotation with Caliph
- Retrieval with Emir
- Prototype Demonstration
- Outlook



Introduction

- Growth of Personal Digital Libraries
 - ◆ 300 Mio. Digital Cameras
 - ◆ 27 Billion Digital Images in 2004
- Emerging Standards
 - ◆ EXIF
 - ◆ MPEG-7
- Representation of Knowledge in Metadata
 - ◆ Annotation
 - ◆ Retrieval



Related Work

Commercial Products

- Google Picasa, Adobe Photoshop Elements, ...
- DAM Products like Virage, Cumulus, Artesia

Research

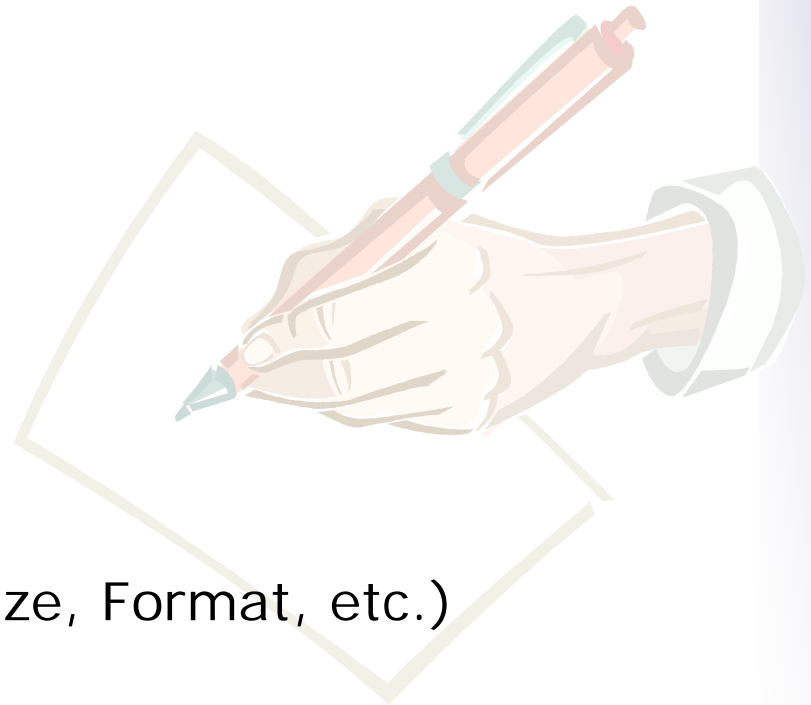
- IBM's Marvel
- VizIR
- Ben Shneiderman's Photofinder
- Know-Center's IMB & Elarm



Annotation with Caliph (1/2)

Extraction of

- MPEG-7 CBIR Descriptors
 - ◆ ColorLayout
 - ◆ ScalableColor
 - ◆ EdgeHistogram
- EXIF and IPTC metadata
- Thumbnail
- Media Instance Metadata (Size, Format, etc.)



Annotation with Caliph (2/2)

Textual Description

- 🌐 Structured (The W's)
- 🌐 Free Text

Quality Rating

- 🌐 Subjective, Scale from 1 to 5

Administrative Metadata

- 🌐 Creator
- 🌐 Meta²



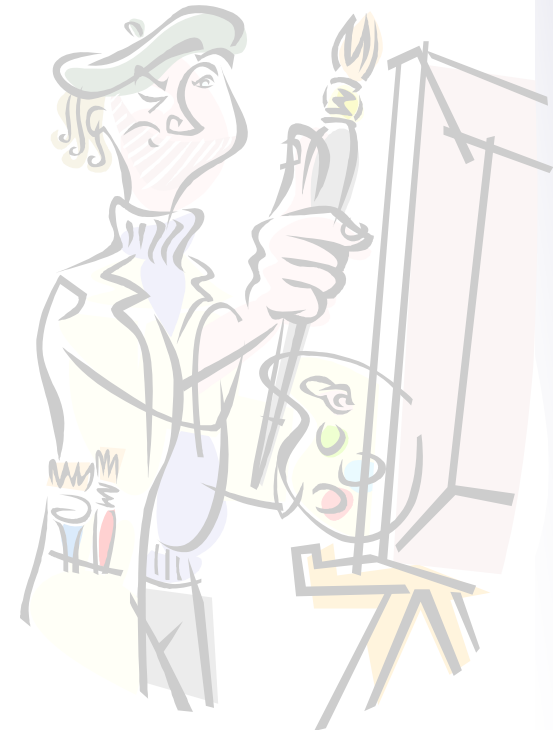
Semantic Annotations (1/3)

Main Goals for Annotation

- Human readable Creation & Presentation
- Support for Computation

Main Tools

- MPEG-7 Descriptors for Semantics
- "Drawing" Annotations
- Library for Semantic Objects



Semantic Annotations (2/3)

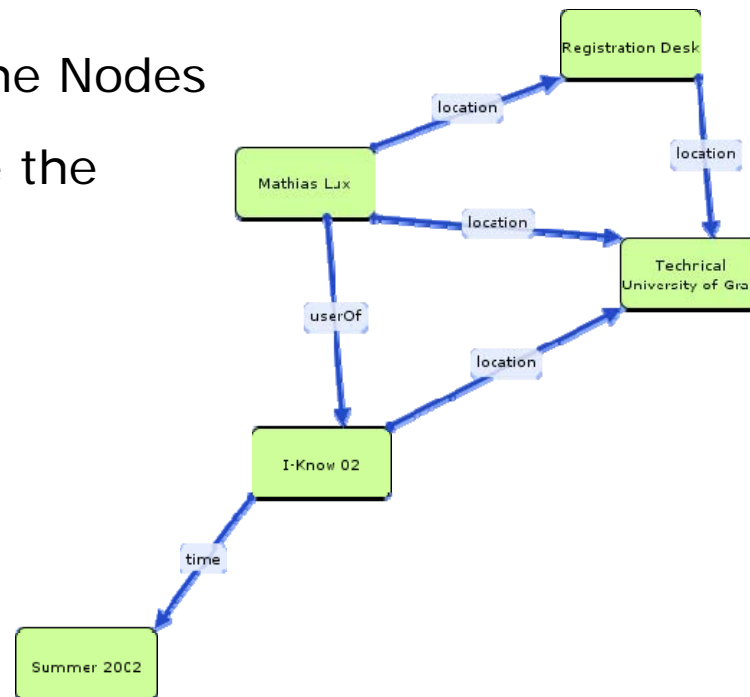
Semantic Annotations are Directed Graphs

Semantic Objects which are the Nodes

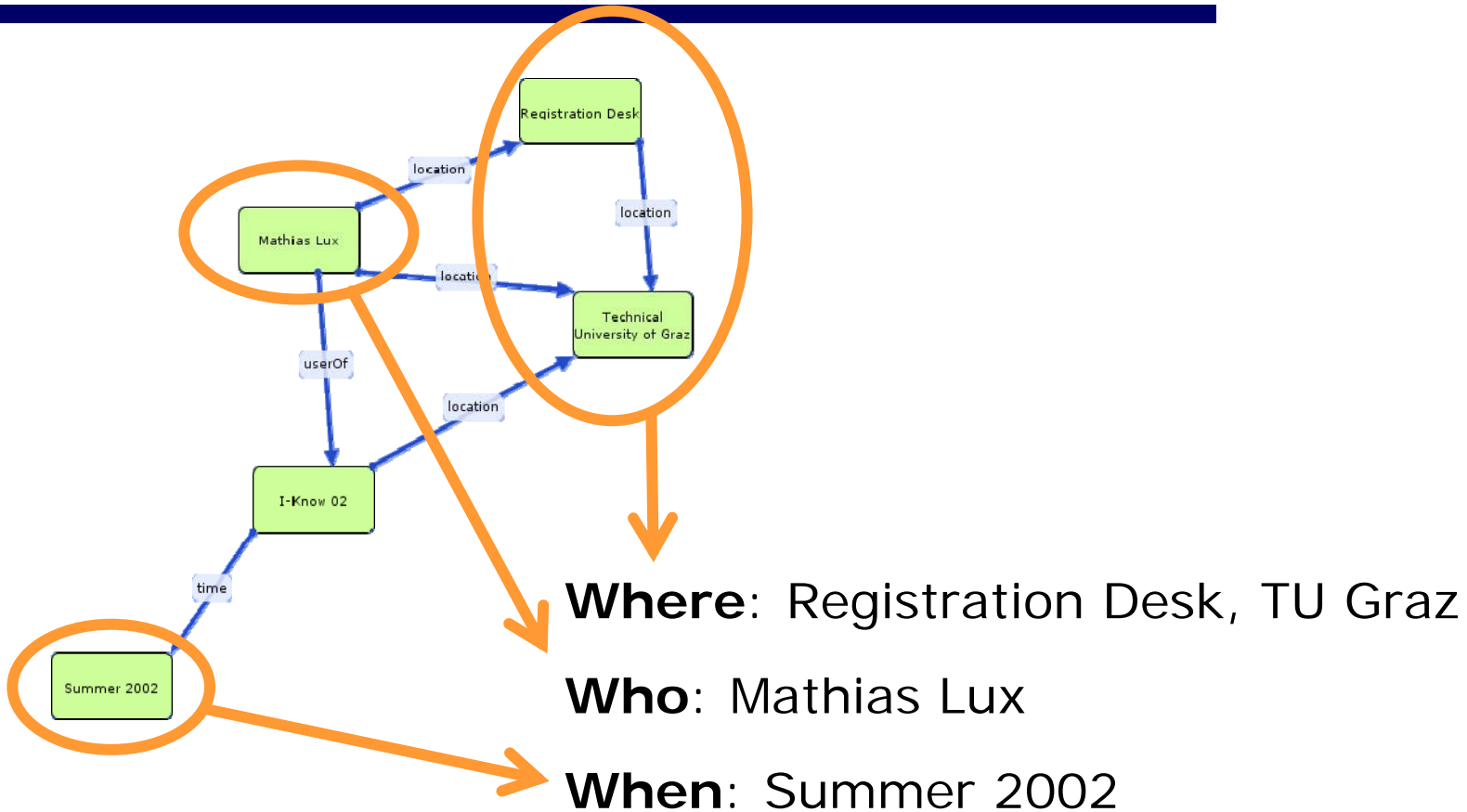
Semantic Relations, which are the edges

Structured Annotation can be extracted

In predefined domains textual descriptions can be generated



Semantic Annotations (3/3)



Annotation Visualization

User defined Graph Layout

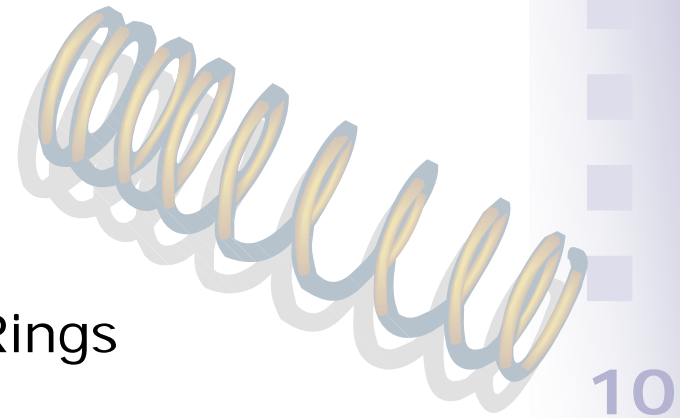
- 🌐 Moving Nodes to desired Position

Radial Layout

- 🌐 Automatic Positioning Nodes around Virtual Center

Spring Embedding

- 🌐 Based on Metaphor:
 - ◆ Node = Metal Ring,
 - ◆ Edge = Spring connecting Metal Rings



Spring Embedding

Algorithm based on Peter Eades' Spring Embedding:

- Springs with logarithmic force: $C1 * \log(d/C2)$
- Nonadjacent nodes repel each other with inverse square law force: $C3/sqr(d)$
- Fixed Number of Iterations:
 - ◆ Calculate Force Vector: f
 - ◆ Move Node along the Force Vector: $C4 * f$

Changes & Additions for Semantic Annotations:

- New Stop Condition: When overall movement "stops"
- Added invisible attracting Centre Node: Separated Graphs are positioned around the Centre



Supported Types of Retrieval:

- Simple Prototype Search Capabilities
 - ◆ Keyword and File Based Retrieval
 - ◆ Content Based Image Retrieval
 - ◆ File Based XPath Search
- Index Based Search Engine
- Search for Semantic Annotations



Semantic Annotation Retrieval

Stable Version supports

- 3 Nodes with 2 Relations
- Wildcards for Nodes and Relations
- Linear Processing Time

Experimental Version supports

- Query Language for Graphs
- Any Number of Nodes and Relations
- Wildcards for Nodes



Demonstration

Emir v 0.8.5 - © 2002-2004 by Mathias Lux, Know-Center

File View Help

Search Search in index XPath search Image search Semantic search Results

Relevance: 0,35
File: [iknow_001.JPG](#) (640 x 480 pixel)
Time: 2002-07-10T18:24:08
Creator: MathiasLux , Eastman Kodak Company DC210 Zoom (V05.00)
Quality rating: n.a.
Description: Michael Granitzer is at Technical University of Graz for I-Know 02
(Description created by: MathiasLux with Caliph v0.9.12)

Relevance: 0,35
File: [iknow_011.JPG](#) (640 x 480 pixel)
Time: 2002-07-11T08:16:30
Creator: MathiasLux , Eastman Kodak Company DC210 Zoom (V05.00)
Quality rating: n.a.
Description: Michael Granitzer is at Technical University of Graz for I-Know 02
(Description created by: MathiasLux with Caliph v0.9.12)

Relevance: 0,35
File: [iknow_042.JPG](#) (480 x 640 pixel)
Time: 2002-07-11T15:17:33
Creator: MathiasLux , Eastman Kodak Company DC210 Zoom (V05.00)
Quality rating: n.a.
Description: Michael Granitzer is at Technical University of Graz for I-Know 02
(Description created by: MathiasLux with Caliph v0.9.12)

Relevance: 0,35

Searched for 2,64 s, formatting lasted 1,08 s

3,5M of 8,1M

by Mathias Lux, Know-Center: iknow_012.JPG

Image Information Semantics Visuals

```

    graph TD
      PeterScheir[Peter Scheir] -- patientOf --> Talking[Talking]
      MarkusStrohmaier[Markus Strohmaier] -- patientOf --> Talking
      Talking -- agentOf --> WalterSarka[Walter Sarka]
      Talking -- location --> TUofGraz[Technical University of Graz]
      TUofGraz -- location --> IKnow02[I-Know 02]
      IKnow02 -- time --> Summer2002[Summer 2002]
      MichaelGranitzer[Michael Granitzer] -- patientOf --> Talking
  
```

Agents

- Andreas Ausserhofer
- Armin Ulbrich
- Bernhard Gissing
- Caroline Gross
- Edi Stoliser
- Erwin Duschning
- Gisela Dösinger
- Herbert Pacnik
- Herwig Rollet
- Ines Puntschart
- Ingrid Fuchs
- Johannes Farmer
- Jutta Becker

Events

- I-Know 02
- I-Know 04
- Presentation
- Reading
- Talking

Places, Times and Objects

- I-Know Marquee
- Know-Center Stand
- Main Hall
- Registration Desk
- Stadthalle Graz
- Summer 2002
- Summer 2004
- Technical University of Graz

Help:
Drag and drop objects from the tables on the right hand side.
left mouse button or *middle mouse button* to draw relations.

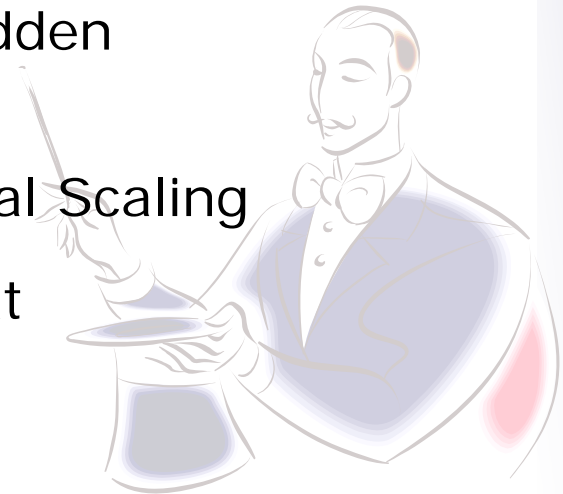
4,9M of 8,0M

Magick is a *Cross Media Retrieval Application*

- Text Documents & Images in Combination are handled independent from their Media Type

Requirements for Magick

- Identification and Visualisation of Hidden Interconnections
- Using Clustering and Multidimensional Scaling
- Equal Treatment for Images and Text



Magick - Concept



Extraction & Normalizing of Metadata

- 🌐 IPTC, Dublin Core, EXIF, HTML Meta Tags

Normalization of the content

- 🌐 ColorLayout, ScalableColor in Case of Images
- 🌐 Words, Sentences, Stemming, Stop Words, ...

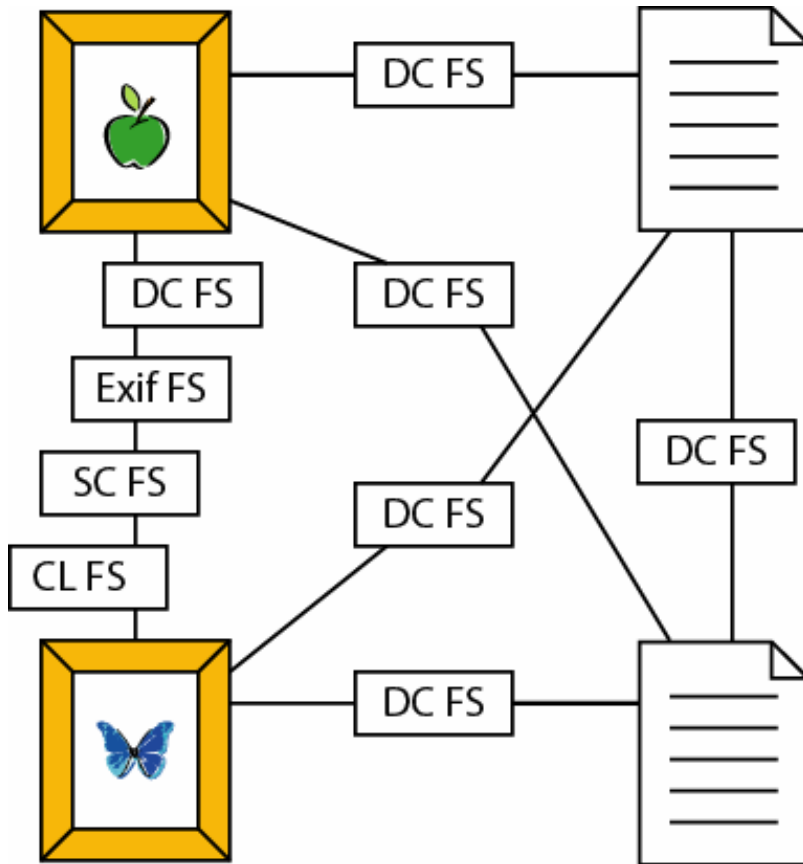
Definition & Linking of Metrics

- 🌐 Similarity and Distance of Content and Metadata

Processing and Visualization

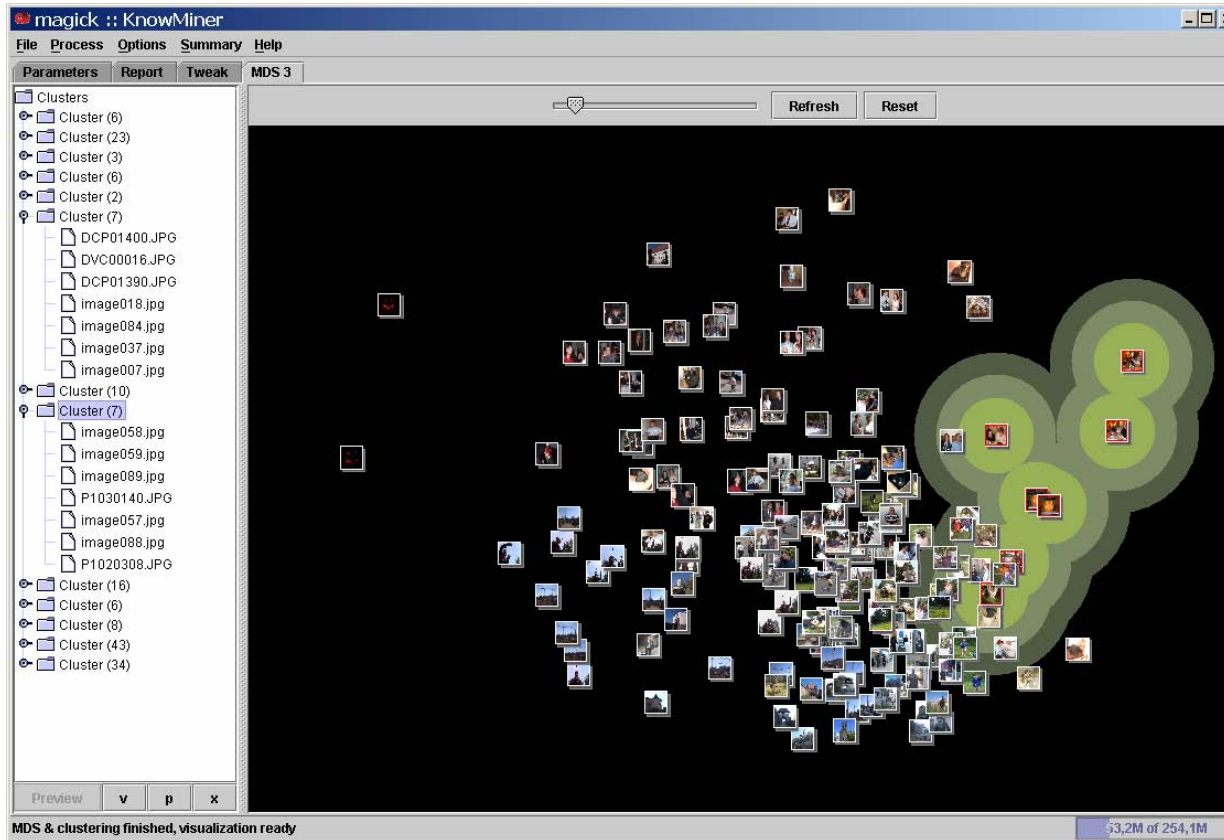
- 🌐 Clustering (HAC) und Multidimensional Scaling (FDP)

Magick – Metrik



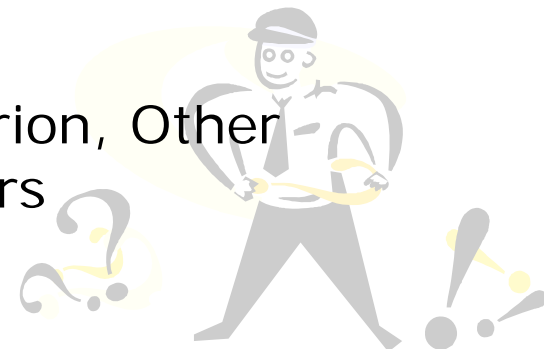
- Feature Spaces
- Weighted Distance Measure scaled to $[0,1]$
- Arithmetic Mean Using Multiple FS

Magick - Visualization



Magick Evaluation

- Implemented Visualization & Retrieval Techniques can be applied in Cross Media Retrieval
- Identification of Main Focuses is Possible
- Documents Need Common Denominator
 - ◆ Otherwise Outliers Will Happen
 - ◆ Common Denominator Can Be Small (e.g. Keywords, Hierarchy)
- Common Denominator is Main Criterion, Other Features Separate too Dense Clusters



Animation: Too Dense Clusters



Future Work

- Integration of Semantic Annotations FS in KnowMiner, especially Magick
- Simplification of Semantic Annotation by
 - ◆ Grouping/Clustering of Images (on EXIF, Timestamps, ...)
 - ◆ Innovative User Interfaces
 - ◆ Automatic Extraction of Concepts (see Marvel)
- Generalization of Semantic Annotations from MPEG-7 to more General Formats (e.g. Domain Ontologies)

