

Multimedia and Semantic Web

Shankar Vembu
(with Stephan Baumann)

Competence Center for Computational Culture
German Research Center for Artificial Intelligence
Kaiserslautern, Germany

Outline

- Semantic web and MPEG-7
- Multimedia ontology
- Semantic analysis
- Ontology learning

Outline

- Semantic web and MPEG-7
- Multimedia ontology
- Semantic analysis
- Ontology learning

Semantic web (in one slide!)

- Annotated resources on the web
- Machine interpretable resources
- Semantic web jargon
 - Resource Description Framework (RDF)
 - RDF Schema, Web Ontology Language (OWL)

MPEG-7

- Metadata standard for multimedia documents
- Comprehensive suite of metadata descriptors
 - Low-level features
 - Structural aspects
 - Semantic aspects
 - Storage format, encoding
 - ...
- Goals
 - Machine interpretable resources
 - Efficient search and retrieval

Problems with MPEG-7

- Emphasis on low-level and structural features
- Not rich in semantics (objects, relationships, events)
- Lack of interoperability with semantic web standards
 - Uses XML Schema as the description language

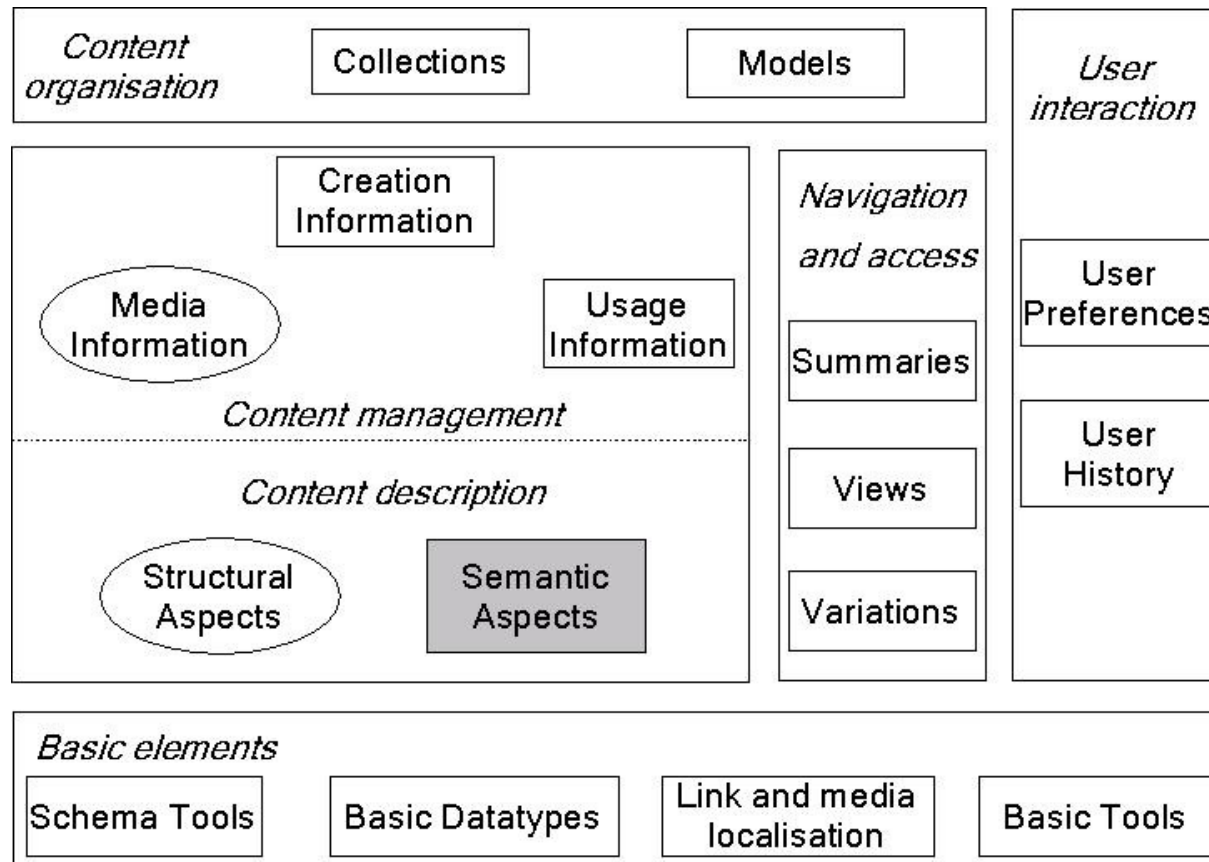
Outline

- Semantic web and MPEG-7
- **Multimedia ontology**
- Semantic analysis
- Ontology learning

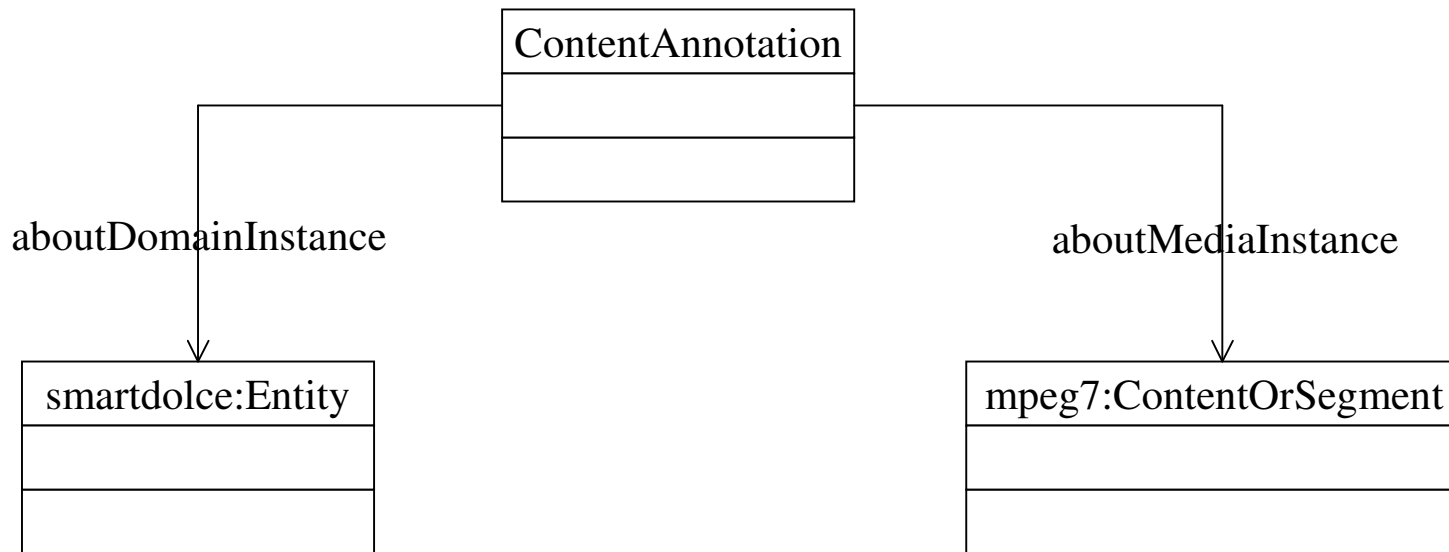
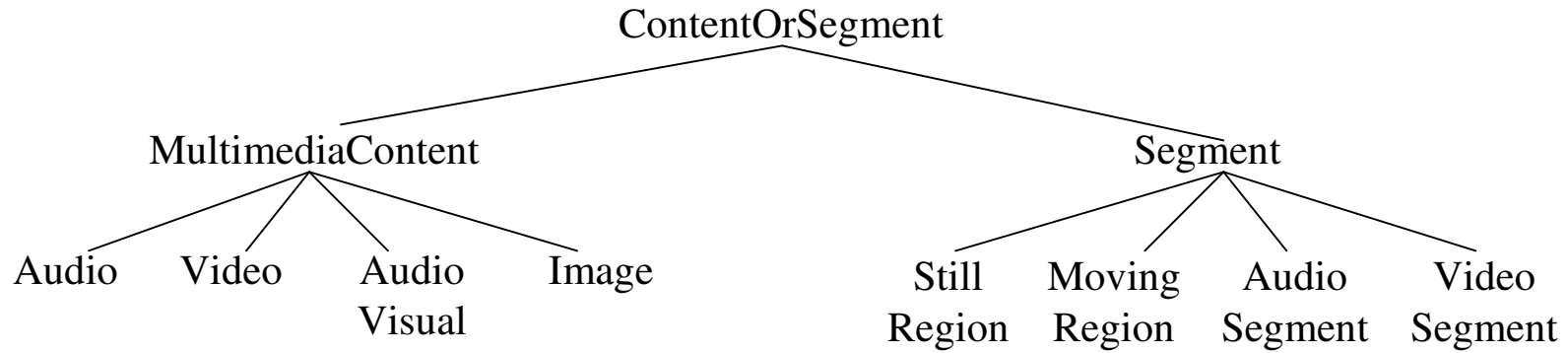
Design approaches (Vembu *et al.* 2006)

- Semantic web + MPEG-7 = Multimedia ontology
- Map MPEG-7 to RDFS or OWL (Hunter 2001)
- (a) Use MPEG-7 to model low-level and structural aspects of multimedia
- (b) Use domain-specific ontologies to model semantics of the domain (eg: sports)
- Merge (a) and (b)

MPEG-7 MDS (taken from the specs)



SmartMedia ontology



Multimedia plugin for Protege

The screenshot displays the Protege 3.0 interface with the multimedia plugin. The main window is titled "mediaexample - Protégé 3.0" and shows a video player displaying a soccer match. The interface is divided into several panels:

- CLASS BROWSER:** Shows a class hierarchy for the project "mediaexample". The "smartmedia" class is highlighted, and its instances are listed in the Instance Browser.
- INSTANCE BROWSER:** Lists instances of the "smartmedia:ContentAnnotation" class, including "Applause for goal", "Assist: Steinar Holvik", "Cross by SCHNEIDER Bernd", "Foot trap by FRINGS Torsten", "goal by BALLACK Michael, 21'", "Goalkeeper Jaromir Blazek", "Pass by SCHWEINSTEIGER", "SCHWEINSTEIGER", and "Shot(on goal) by BALLACK Michael".
- Media Player:** Displays a video of a soccer match. A callout points to this area with the label "Media player".
- List of annotations:** A table listing annotations with their descriptions and snapshots. A callout points to this area with the label "List of annotations".
- INSTANCE EDITOR:** Shows the details for the instance "Shot(on goal) by BALLACK Michael". A callout points to this area with the label "Media instance".
- SmartMedia ontology:** A callout points to the "smartmedia:ContentAnnotation" class in the Instance Browser with the label "SmartMedia ontology".

Description	Snapshot
Goalkeeper Jaromir Blazek smartmedia:ContentAnnotation 0:0:2:20	
Foot trap by FRINGS Torsten smartmedia:ContentAnnotation 0:0:5:36	
Cross by SCHNEIDER Bernd smartmedia:ContentAnnotation 0:0:13:48	
Shot(on goal) by BALLACK Michael smartmedia:ContentAnnotation 0:0:16:8	
goal by BALLACK Michael, 21' smartmedia:ContentAnnotation 0:0:17:58	

What next?

- Multimedia ontologies are in place
- Annotation of multimedia content is possible
- (semantic) Web search is possible
 - Text-based queries
- BUT...

What next?

- Who annotates the content?
 - Manual annotation ☹️
- Need to (semi)automate the annotation process
 - Using machine learning approaches
 - Using content-based retrieval techniques

Outline

- Semantic web and MPEG-7
- Multimedia ontology
- Semantic analysis
- Ontology learning

Semantic analysis

- Annotation with low-level MPEG-7 features is possible
- Annotation at a semantic level of abstraction is non-trivial
 - Identify concepts in images
 - Identify events in videos

Probabilistic models (Naphade *et al.* 1998)

- **Multijects:** Probabilistic multimedia objects
 - Visual concepts: *sky, water, forest* etc.
 - Audio-visual concepts: *explosion, waterfall*
- **Multinet:** Bayesian network of multijects
 - Models the relationship between concepts

Probabilistic models (Barnard *et al.* 2003)

- Probabilistic models of text and images
 - Uses captions found in image databases
- Uses multi-modal extensions of LDA
 - Latent Dirichlet Allocation: Topic model for semantic analysis of text documents
- Text-based annotation and *querying* is now possible

Concept classifiers (Vogel and Schiele 2006)

- Classifies local image regions into concepts
 - *sky, rocks, flowers* etc.
- No segmentation step
 - Uses a 10*10 grid on the image
 - Classifies each grid point into a concept class
 - Concept-occurrence vector
- Notion of semantic similarity

Outline

- Semantic web and MPEG-7
- Multimedia ontology
- Semantic analysis
- **Ontology learning**

Ontology learning

- Why learn ontologies?
 - Knowledge acquisition bottleneck
- Data-driven acquisition of knowledge
- Ontology learning from text
 - Abundance of text data on the web
 - Uses NLP techniques
 - (Tutorial at ECML/PKDD 2005)

Challenges, Open questions

- Multimedia ontology population
- Multimedia ontology learning (Karkaletsis *et al.* 2005)
- To what extent is it possible to extract metadata descriptors from multimedia documents so as to enable their subsequent search and retrieval using semantic web technologies?

Acknowledgements

- SmartWeb project
 - <http://www.smartweb-project.org>
- Michael Sintek (DFKI GmbH)
- Malte Kiesel (DFKI GmbH)

References

- K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. Blei, M. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3:1107-1135, February 2003.
- J. Hunter. Adding multimedia to the semantic web - Building an MPEG-7 ontology. In *Proceedings of the International Semantic Web Working Symposium*, Stanford University, California, USA, July 30 - August 1 2001.
- ISO/IEC 15938-5. FCD Information technology - Multimedia content description interface - Part 5: Multimedia description schemes, 2003.
- V. Karkaletsis, G. Paliouras, C. D. Spyropoulos. A bootstrapping approach to knowledge acquisition from multimedia content with ontology evolution. In *Proceedings of the International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning*, pp. 98-105, Helsinki University of Technology, Finland, June 2005.
- M. Naphade, T. Kristjansson, B. Frey, T.S. Huang. Probabilistic multimedia objects (multijects): A novel approach to indexing and retrieval in multimedia systems. In *Proceedings of the 5th IEEE International Conference on Image Processing*, Chicago, IL, Oct 1998.
- S. Vembu, M. Kiesel, M. Sintek, S. Baumann. Towards bridging the semantic gap in multimedia annotation and retrieval. In *Proceedings of the 1st International Workshop on Semantic Web Annotations for Multimedia at the 15th International World Wide Web Conference*, Edinburgh, Scotland, May 2006. (to appear)
- J. Vogel, B. Schiele. Semantic Scene Modeling and Retrieval for Content-Based Image Retrieval. Accepted for publication at *International Journal of Computer Vision*, 2006.

Are we ready to embrace the
semantic web?