Semantic Web and Multimedia

- Multimedia Ontology -

Steffen Staab
http://isweb.uni-koblenz.de
"One Ring to rule them all,  
One Ring to find them,  
One Ring to bring them all and  
in the darkness bind them."

*Inscribed on the One Ring*

For ~99% of multimedia people the answer for content annotation is …

**MPEG-7!**

MPEG-7 has standardized more than 140 classification schemes that describe properties of multimedia content.

BUT: what did we learn from eCommerce?

→ An XML standard is per se not the solution for a general information integration problem!
Meaning of Informationen:
(or: what it means to be a computer)
XML ≠ Meaning, XML = Structure

ISWeb - Information Systems & Semantic Web

Steffen Staab
staab@uni-koblenz.de
What is the Problem with MPEG-7?

How do you formulate a query to get all segments that show „Sky“?

First Shot:

XQL: //StillRegion[.//Keyword="Sky"]
<Mpeg7>
  <Description xsi:type="ContentElementType">
    <MultimediaContent xsi:type="ImageType">
      <Image id="IM01">
        <SpatialDecomposition>
          <StillRegion id="SR1">
            <Semantic>
              <Label><Name>Roosevelt</Name></Label>
            </Semantic>
          </StillRegion>
          <StillRegion id="SR2">
            <TextAnnotation>
              <KeywordAnnotation><Keyword>Churchill</Keyword></KeywordAnnotation>
            </TextAnnotation>
          </StillRegion>
          <StillRegion id="SR3">
            <Semantic>
              <Definition>
                <StructuredAnnotation><Who><Name>Stalin</Name></Who></StructuredAnnotation>
              </Definition>
            </Semantic>
          </StillRegion>
        </SpatialDecomposition>
      </Image>
    </MultimediaContent>
  </Description>
</Mpeg7>
Scenario

A history of G8 violence (video) (© Reuters)

Localize a segment
Annotate the content

G8 Summit, Heiligendamm, 2007 EU Summit, Gothenburg, 2001

Link to the knowledge on the Web
:Seq1 foaf:depicts dbpedia:34th_G8_Summit
:Seq4 foaf:depicts dbpedia:EU_Summit
geo:Heiligendamm skos:broader geo:Germany
MPEG-7 and the Semantic Web

MDS Upper Layer represented in RDFS
- 2001: Hunter
- Later on: link to the ABC upper ontology

MDS fully represented in OWL-DL
- 2004: Tsinaraki et al., DS-MIRF model

MPEG-7 fully represented in OWL-DL
- 2005: Garcia and Celma, Rhizomik model
- Fully automatic translation of the whole standard

MDS and Visual parts represented in OWL-DL
- 2007: Arndt et al., COMM model
- Re-engineering MPEG-7 using DOLCE design patterns
Requirements [aceMedia, MMSEM XG]

MPEG-7 compliance
  - Support most descriptors (decomposition, visual, audio)

Syntactic and Semantic interoperability
  - Shared and formal semantics represented in a Web language (OWL, RDF/XML, RDFa, etc.)

Separation of concerns
  - Domain knowledge versus multimedia specific information

Modularity
  - Enable customization of multimedia ontology

Extensibility
  - Enable inclusion of further descriptors (non MPEG-7)
# MPEG-7 Based Ontologies

<table>
<thead>
<tr>
<th>Foundational Ontologies</th>
<th>Hunter</th>
<th>DS-MIRF</th>
<th>Rhizomik</th>
<th>COMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>None</td>
<td>None</td>
<td>DOLCE</td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>OWL-Full</td>
<td>OWL-DL</td>
<td>OWL-DL</td>
<td>OWL-DL</td>
</tr>
<tr>
<td>Coverage</td>
<td>MDS+Visual</td>
<td>MDS+CS</td>
<td>All</td>
<td>MDS+Visual</td>
</tr>
<tr>
<td>Applications</td>
<td>Digital Libraries</td>
<td>Digital Libraries</td>
<td>Digital Rights</td>
<td>MM Analysis</td>
</tr>
</tbody>
</table>
COMM by simple examples
Example 1: Fragment Identification


dns:realized-by

core:image-data

dns:plays

dns:played-by

dns:defines

loc:region-locator-descriptor

loc:spatial-mask-role

loc:bounding-box

5 25 10 20 15 15 10 10 5 15"^^xsd:string
Example 1: Region Annotation


core:image-data

dns:setting

dns:realized-by

dns:plays

dns:played-by

dns:defines

dns:played-by

loc:region-locator-descriptor

loc:bounding-box

data:has-rectangle

loc:spatial-mask-role

core:semantic-annotation

core:semantic-label-role

foaf:Person

5 25 10 20 15 15 10 10 5 15"^^xsd:string
Example 2: Fragment Identification


dns:realized-by

core:video-data
dns:plays

dns:played-by

dns:defines

loc:media-time-descriptor

loc:temporal-mask-role
dns:played-by

loc:media-time-point

data:has-time

"1:21"^^xsd:time
Example 2: Sequence Annotation


core:video-data

dns:setting
dns:played-by
dns:realized-by

dns:defines

dns:played-by

geo:Sweden

skos:broader

loc:temporal-mask-role

core:semantic-label-role

dns:defines

dns:played-by

geo:Gothenburg

"1:21"^^xsd:time

data:has-time

loc:media-time-point

loc:media-time-descriptor

dns:defines

dns:played-by

ISWeb - Information Systems & Semantic Web
Steffen Staab
staab@uni-koblenz.de
Implementation

COMM fully formalized in OWL DL

- Rich axiomatization, consistency check (Fact++v1.1.5)
- OWL 1.1: qualified cardinality restrictions for number restrictions of MPEG-7 low-level descriptors

JAVA API available

- MPEG-7 class interface for the construction of meta-data at runtime
…now some of the interiors…
COMM: Modules

Annotation Pattern

Decomposition Pattern
Multimedia ontology consists of

- **Core module** that contains the design patterns

- Modules that specialize the core module for **different media types**

- Modules that contain **media independent MPEG-7 description tools** such as media information or creation & production

- **Data type module** that formalizes MPEG-7 data types e.g. matrices, vectors, unsigned-int-5, float-vector, probability-vector, …
DOLCE Foundational Ontology

4D world view centered around

- **Endurants**: Independent wholes that exist in time and space
- **Perdurants**: Events, processes, phenomena, …

DOLCE is a library of foundational ontologies that provides 2 design patterns (extensions) that are especially important for MPEG-7:

- **Ontology of Information objects (OIO)**: Formalization of information exchange
- **Descriptions & Situations (D&S)**: Formalization of context

Use these extensions to translate the technical concepts of MPEG-7 in the DOLCE vocabulary
DOLCE Foundational Ontology
Ontology of Information Objects (OIO)

Formalization of information exchange
- Shannon’s communication theory
- Communication elements by Jakobson

Information object represents pure abstract information (message)

Relevance for multimedia ontology:
- MPEG-7 describes digital data (multimedia information objects) with digital data (annotation)
- Digital data entities are information objects
Ontology of Information Objects

A - Algorithm & Digital Data
Descriptions & Situations (D&S)

Distinction between:
- DOLCE ground entities (regions, endurants, perdurants)
- Descriptive entities (parameters, roles, courses)

Descriptions
- Formalize context
- Define descriptive concepts

Situations
- Are explained by descriptions
- Are settings for ground entities

Don’t confuse a situation and its description. The situation is unique, its descriptions may be conflicting!
Relevance for multimedia ontology:
- Meaning of digital data depends on context
- Digital data entities are connected through computational situations (e.g. input and output data of an algorithm)
- Algorithms are descriptions
- Annotations and decompositions are situations that satisfy the rules of an algorithm / method
Example

Information Object „Secure the building“
Information Realization
Information Encoding: ASCII-Code decimal
About: the White House
Situation: Securing the president
Expresses:
  1. Buy the building
  2. Everyone out of the building, blinds shut down
  3. Bomb the building
Decomposition Pattern

ISWeb - Information Systems & Semantic Web
Steffen Staab
staab@uni-koblenz.de
Locators

Need to be specialized for different media
An annotation is subjective – due to algorithm or manual process
A Semantic annotation links to an outside ontology („Particular“)
The patterns cover more than just MPEG-7, e.g. other document formats.
Decomposing image data

G – Decomposition of Images
KAT Annotation Tool

ISWeb - Information Systems
Steffen Staab
staab@uni-koblenz.de
Multimedia Data

ISWeb - Information Systems & Semantic Web

Steffen Staab
staab@uni-koblenz.de

Multimedia Data
### Modeling Decisions:

- DS-MIRF and Rhizomik: 1-to-1 translation from MPEG-7 to OWL/RDF
- Hunter: Simplification and link to the ABC upper model
- COMM: NO 1-to-1 translation
  - Need for patterns: use DOLCE, a well designed foundational ontology as a modeling basis

### Scalability (cf Troncy et al 2007):

<table>
<thead>
<tr>
<th></th>
<th>Hunter</th>
<th>DS-MIRF</th>
<th>Rhizomik</th>
<th>COMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triples</td>
<td>11</td>
<td>27</td>
<td>20</td>
<td>19</td>
</tr>
</tbody>
</table>
Example Query

```
SELECT ?ITEM ?URI ?TYPE
WHERE {
  ?ITEM custom:directInstanceOf ?TYPE;
  a core:multimedia-data;
  core:plays core:root-segment-role;
  core:realized-by ?URI;
  ?annotation1 a core:semantic-annotation;
  core:setting-for ?ITEM;
  core:setting-for ?label1;
  core:satisfies [
    a core:method;
    core:defines ?annotated-data-role1;
    core:defines ?semantic-label-role1].
  ?semantic-label-role1 a core:semantic-label-role.
  ?annotation2 a core:semantic-annotation;
  core:setting-for ?ITEM;
  core:setting-for ?label2;
  core:satisfies [
    a core:method;
    core:defines ?annotated-data-role2;
    core:defines ?semantic-label-role2].
  ?semantic-label-role2 a core:semantic-label-role.
  ?label1 a ex:USLeader.
  ?label2 a ex:USSRLeader.
}
```
Implications

Challenges on data management:

- Many triples for one annotation
  - But: can hardly be avoided with standard notations
  - Data partitioning required – and possible
  - On demand access required (hibernation)

- Complex queries

- Path queries

- Recursive queries desired (e.g. using networked graphs)

- Linking to other data sets
Capabilities and Maturity Levels

Former Situation:
no standard, no vocabulary, no coupling of data and application unless by tiresome 1:1 agreement of involved parties

Current situation:
MPEG-7, tight coupling, standard vocabulary, agreement between involved parties on which vocabulary to use and its exact meaning

Future / desired situation:
loose coupling, standard vocabulary with pre-defined meaning, automatic ad-hoc coupling of data and integration