Connecting Foundational Ontologies with MPEG-7 Ontologies for Multimodal QA
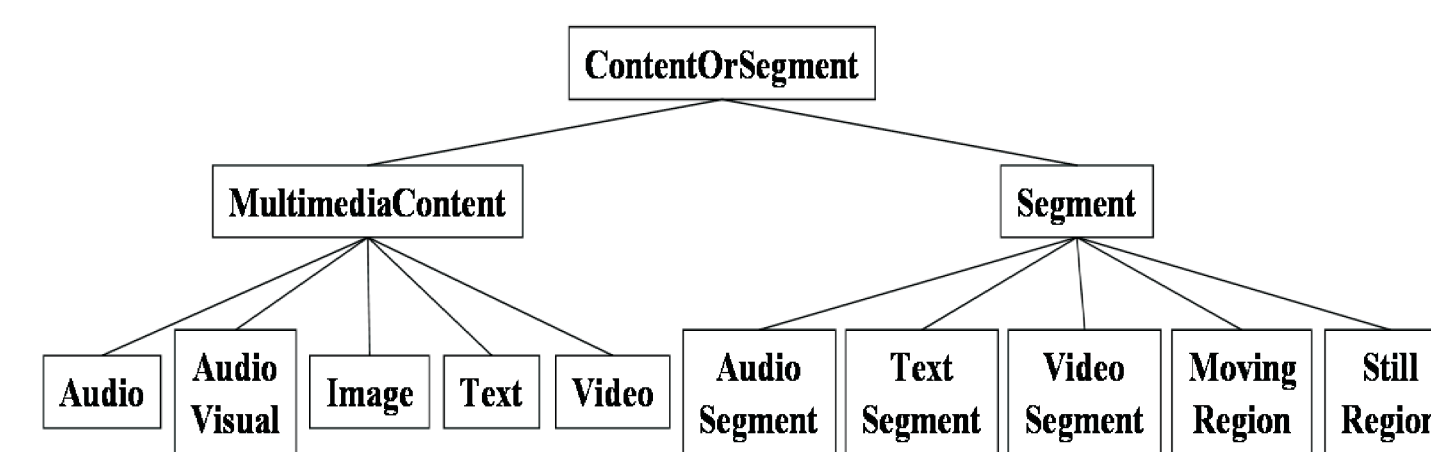
Massimo Romanelli, Daniel Sonntag, Norbert Reithinger {massimo.romanelli,daniel.sonntag,norbert.reithinger}@dfki.de

DFKI GmbH, Saarbrücken, Germany

Smartweb Requirements

- Multimodal dialogue with question answering on PDA.

- Multimodal recognition for speech and gestures.

- Multimodal interaction with different kinds of result.

- Modality interpretation, fusion, and intention processing.

- Modality fission, result rendering for text, images, videos, graphics, and synthesis of speech.

- Multimedia presentation and interaction.

- Ontology Framework for multimodal multimedia interaction.

- Clear, ontology-based interface between modules (Jena API).

- Real-time interactive editing of semantic queries.

Ontology Architecture

- The SWIntO (SmartWeb Integrated Ontology), is based on a upper model ontology realized by merging well chosen concepts from two established foundational ontologies, DOLCE and SUMO.

- The domain-specific knowledge like sport-event, navigation, or webcam is defined in dedicated ontologies.

- Multimodal interaction on multimedia results is based on the Discourse and the Smartmedia ontology.

Smartmedia

- Link to Upper Model Ontology (smartmedia:aboutDomainInstance).

- Multimedia Decomposition (mpeg7:SegmentDecomposition) in space, time, frequency.

- File format and coding parameters (mpeg7:MediaFormat).

- MPEG7-based media annotation.

Smartmedia Connection
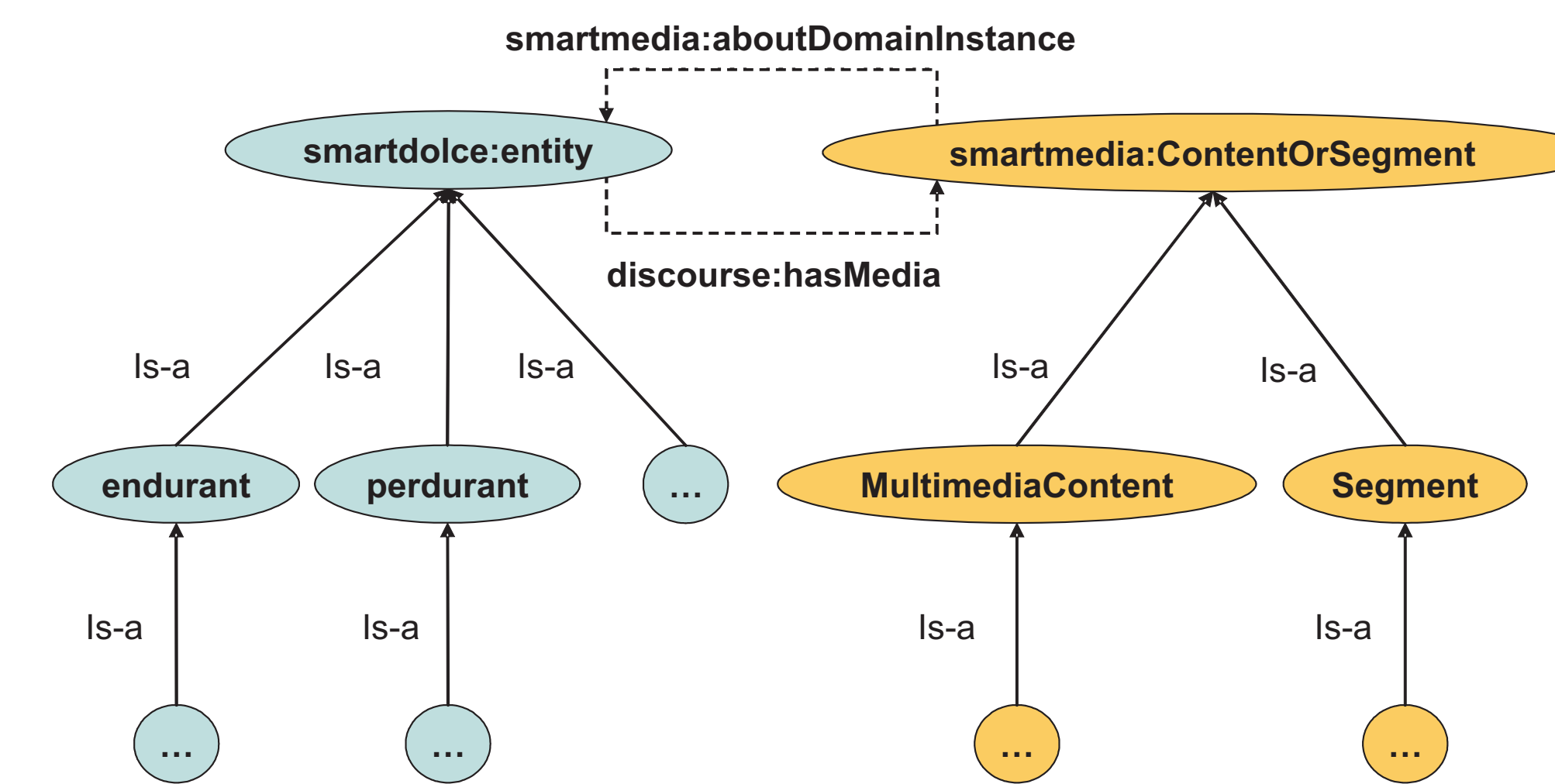
- The surface level referring to the properties of realized media as in the Smartmedia and the deep semantic representation of these objects are linked with the **smartmedia:aboutDomainInstance** property. In this way the link to the upper model ontology is inherited to all segments of a media instance decomposition.

- Through the **discourse:hasMedia** property located in the smartdolce:entity top level class and inherited to each concept in the ontology we realize a pointer back to the Smartmedia ontology.
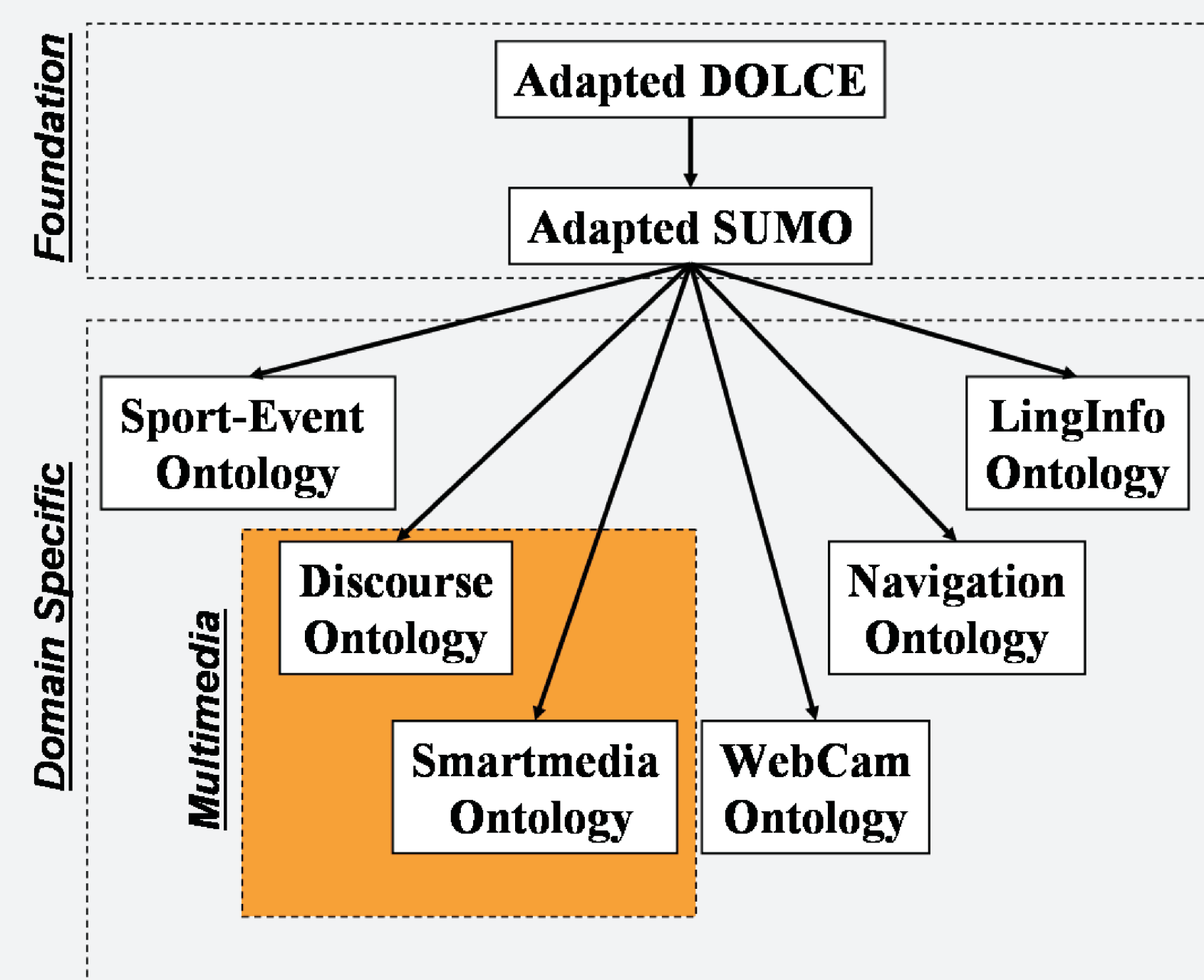
RDF Example

```
<emma:container>
  <emma:Interpretation rdf:about="http://smartweb.semanticweb.org/ontology/emma#GestureInterpretation">
    <emma:applicationInstanceData>
      <discourse:PointingGesture rdf:about="http://smartweb.semanticweb.org/ontology/emma#PGesture_1">
        <discourse:objectReference>
          <sportevent:GoalKeeper rdf:about="http://smartweb.semanticweb.org/ontology/sportevent#swinto_Instance_5">
          ...
          </sportevent:GoalKeeper>
        </discourse:objectReference>
        <discourse:objectReference rdf:reference="http://smartweb.semanticweb.org/ontology/mpeg7#Taffarel"/>
        <discourse:objectReference>
          <mpeg7:Image rdf:about="http://smartweb.semanticweb.org/ontology/mpeg7#i18811">
            <mpeg7:stillRegion>
              <mpeg7:StillRegion rdf:about="http://smartweb.semanticweb.org/ontology/mpeg7#Brasil">
                <mpeg7:aboutDomainInstance rdf:reference="http://smartweb.semanticweb.org/ontology/sportevent#swinto_Instance_2"/>
                <mpeg7:spatialDecomposition>
                  <mpeg7:StillRegionSpatialDecomposition rdf:about="http://smartweb.semanticweb.org/ontology/mpeg7#Decomposition42">
                    <mpeg7:stillRegion>
                      <mpeg7:StillRegion rdf:about="http://smartweb.semanticweb.org/ontology/mpeg7#Taffarel">
                        <mpeg7:aboutDomainInstance rdf:reference="http://smartweb.semanticweb.org/ontology/sportevent#swinto_Instance_5"/>
                      </mpeg7:StillRegion>
                    </mpeg7:stillRegion>
                    <mpeg7:stillRegion>
                      <mpeg7:StillRegion rdf:about="http://smartweb.semanticweb.org/ontology/mpeg7#Ronaldo">
                        <mpeg7:aboutDomainInstance rdf:reference="http://smartweb.semanticweb.org/ontology/sportevent#swinto_Instance_7"/>
                      </mpeg7:StillRegion>
                    </mpeg7:stillRegion>
                    <mpeg7:stillRegion>
                      <mpeg7:StillRegion rdf:about="http://smartweb.semanticweb.org/ontology/mpeg7#Rivaldo">
                        <mpeg7:aboutDomainInstance rdf:reference="http://smartweb.semanticweb.org/ontology/sportevent#swinto_Instance_9"/>
                      </mpeg7:StillRegion>
                    </mpeg7:stillRegion>
                    ...
                  </mpeg7:StillRegionSpatialDecomposition>
                </mpeg7:spatialDecomposition>
              </mpeg7:StillRegion>
            </mpeg7:stillRegion>
          </mpeg7:Image>
        </discourse:objectReference>
      </discourse:PointingGesture>
    </emma:applicationInstanceData>
  </emma:Interpretation>
</emma:container>
```

Standards

- W3C EMMA: for multimodal I/O representation, time stamps, status information, process numbers, generic container for content structures.

- SmartWeb SWEMMA: EMMA Extension for representing queries, results, and status objects.

- RDF for interaction with Semantic Mediator.

- MPEG-7 (ISO/IEC 15938): standard for describing Multimedia content data.