Ontology for Multimedia Applications

Hiranmay Ghosh
TCS Innovation Labs, Delhi

Contributors

Santanu Chaudhury
IIT Delhi

Anupama Mallik
IIT Delhi

Hiranmay Ghosh
TCS / IIT Delhi

... and many other students

Agenda

Part I

- Introduction
- Semantic Web and Ontology
- Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

Part II

- Multimedia Web Ontology Language
- Application Examples
- Distributed Multimedia Applications
- Conclusion

Part I

Agenda

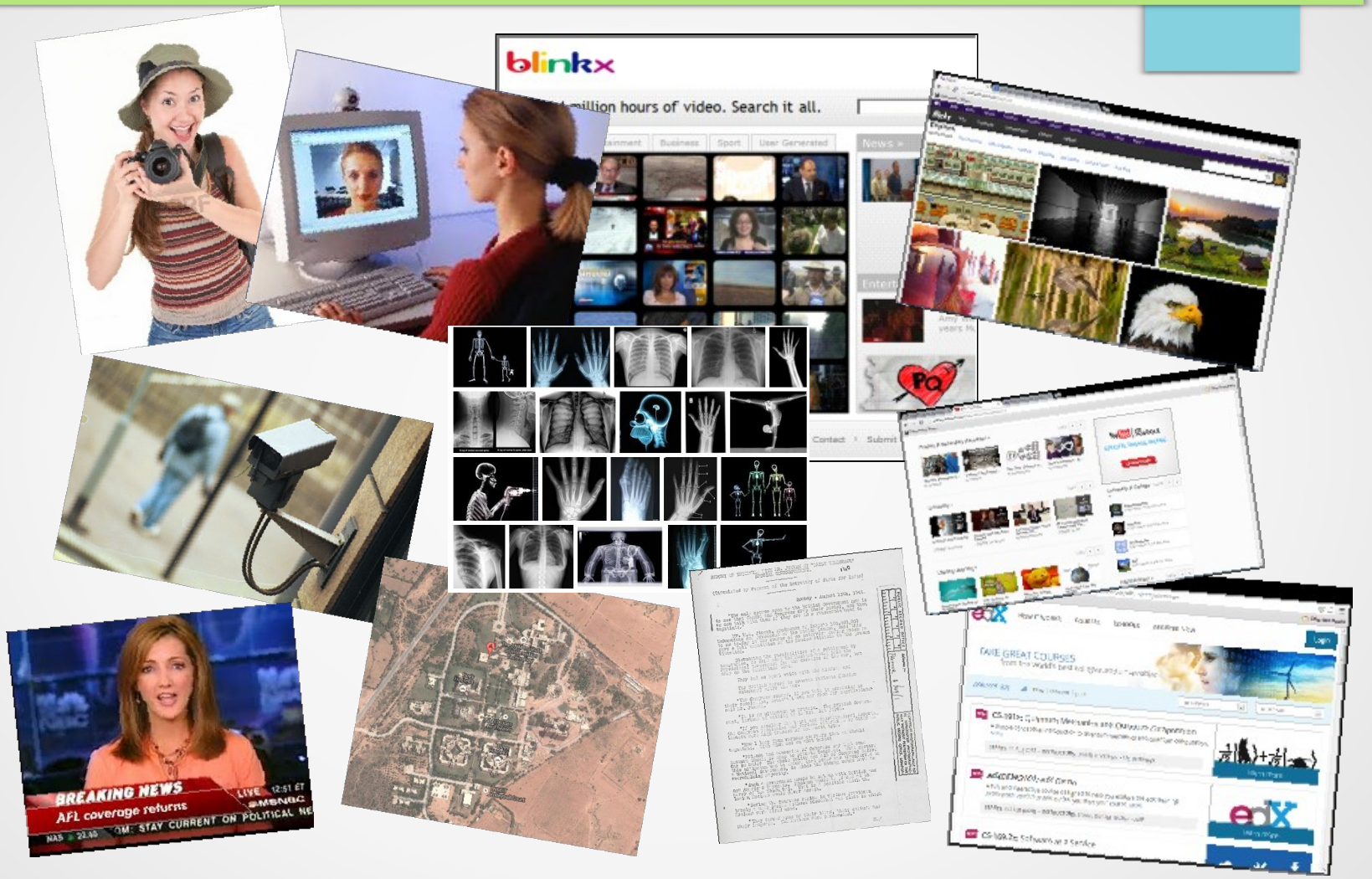
Part I

- **Introduction**
- Semantic Web and Ontology
- Semantic Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

Part II

- Multimedia Web Ontology Language
- Application Examples
- Distributed Multimedia Applications
- Conclusion

Multimedia for infotainment



Some statistics [2012]

Source: pingdom.com

Video

- **14 million** – Number of Vimeo users.
- **200 petabytes** – Amount of video played on Vimeo during 2012.
- **150,648,303** – Number of unique visitors for video to Google Sites, the number one video property (September).
- **1 billion** – PSY's Gangnam Style video became the first online video to reach 1 billion views (currently just over 1.1 billion) and it achieved it in just 5 months.
- **2.7 billion** – Number of views of videos uploaded to YouTube tagged Obama or Romney during the 2012 U.S. election cycle
- **2.5 million** – Number of hours of news-related video that was uploaded to YouTube.
- **8 million** – The number of concurrent viewers of the life Baumgartner's jump from the edge of space, the most e
- **4 billion** – Number of hours of video we watched on Yo
- **60 million** – Number of global viewers monthly on Ustre
- **16.8 million** – Number of total viewers in a 24 hour peri the most ever.
- **181.7 million** – Number of total unique viewers of onlin during December.

Images

- **7 petabytes** – How much photo content Facebook added every month.
- **300 million** – Number of new photos added every day to Facebook.
- **5 billion** – The total number of photos uploaded to Instagram since its start, reached in September 2012.
- **58** – Number of photos uploaded every second to Instagram.
- **1** – Apple iPhone 4S was the most popular camera on Flickr.

How do we

Deal effectively with the large volume of distributed multimedia data?

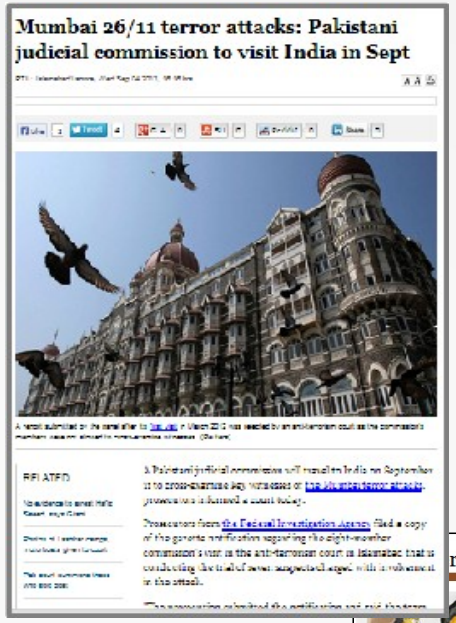
Organize

Retrieve

Navigate

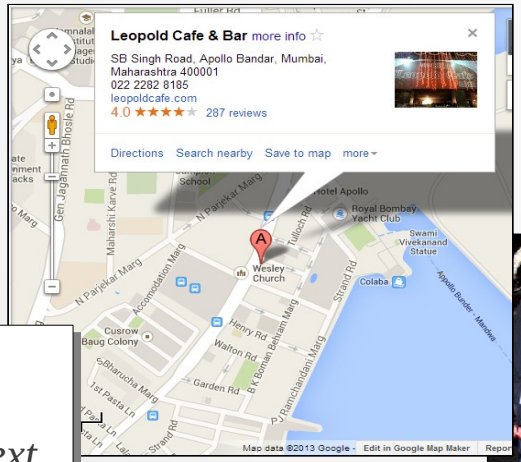
Correlate

News aggregation



- *TV Channels*
- *Newspapers*
- *Social Media*
- *Maps*

- *Aggregate*
- *Summarize*
- *Present*
- *Navigate*



- *Speech*
- *Video*
- *Overlay Text*
- *Image*
- *Text*

Digital Heritage

- *Dance forms*
- *Music genres*
- *Instruments*
- *Myth*
- *Scripture*
- *Artistes*
- *Schools …*

- *Videos*
- *Still images*
- *Document images*
- *Text*

- *Retrieve*
- *Navigate*

Agenda

Part I

- Introduction
- **Semantic Web and Ontology**
- Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

Part II

- Multimedia Web Ontology Language
- Application Examples
- Distributed Multimedia Applications
- Conclusion

The Semantic Web

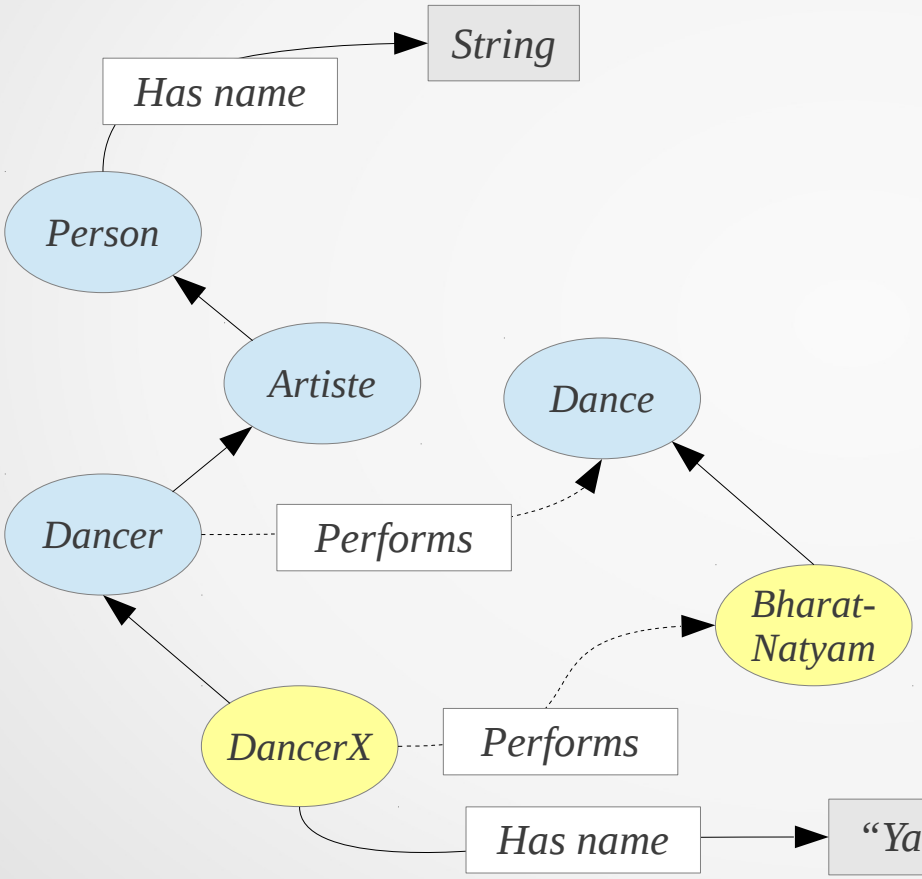
- Semantic data modeling
  - Concepts represented through symbols
  - Relations between the concepts

- Common reference for interpretation of data from multiple sources
- Layers for
  - Syntactic compatibility (XML)
  - Semantic interoperability (RDF, OWL)

W3C Standards

Ontology

- A formal representation of a domain



- *An artiste is a person*
- *A person has name [string]*
- *A dancer is an artiste*
- *A dancer performs dance*
- *DancerX is a dancer*
- *Bharatnatyam is a Dance*
- *DancerX performs Bharatnatyam*
- *DancerX has name “Yamini Krishnamurthy”*

Why use ontology?

- Template for information extraction

  <dancer><name><dance-type>

- Reasoning to find new facts (not explicitly stated)
  - *DancerX* is a *person*
  - *DancerX* performs *Dance*
  - *At least one dancer performs Bharatnatyam*

- Separation of knowledge from program logic facilitates
  - *Knowledge Engineering*
  - *Reuse and maintenance*

Agenda

Part I

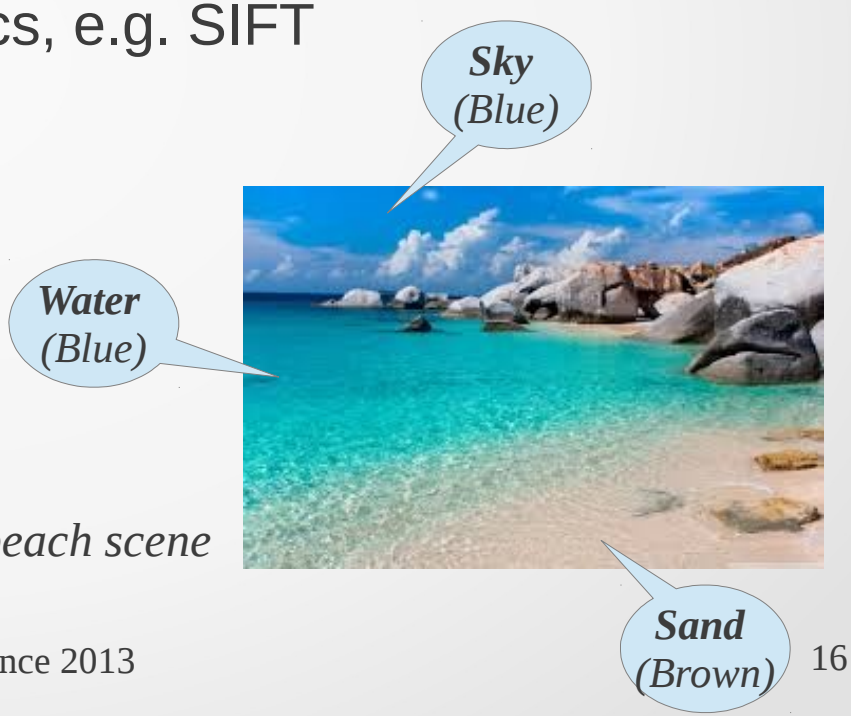
- Introduction
- Semantic Web and Ontology
- **Multimedia Content Processing**
- Ontology for Multimedia Data Interpretation

Part II

- Multimedia Web Ontology Language
- Application Examples
- Distributed Multimedia Applications
- Conclusion

Content, Concept & Context

- Content based retrieval (early 1990's)
  - *Low level image features, e.g. Color & texture*
- Concept based (late 1990's – still evolving)
  - Features conveying more semantics, e.g. SIFT
  - Machine Learning techniques
- Contextual reasoning
- Granularity of semantics
  - Scene recognition
  - Object recognition
    - *Generic & Specific*



A beach scene

Current state of content understanding

- Significant progress in visual data understanding
  - *Document images, Surveillance, Medical / Satellite imagery, Scene understanding, Action recognition, ...*
- Audio & Speech
  - *Good progress*
- Domain specific solutions
  - *Implicit domain knowledge*

Semantic gap: still an unsolved problem

Semantic World

STOP !

Joy and freedom

Bananas

Red Light

Bharatnatyam

Semantic Gap

Media World

Agenda

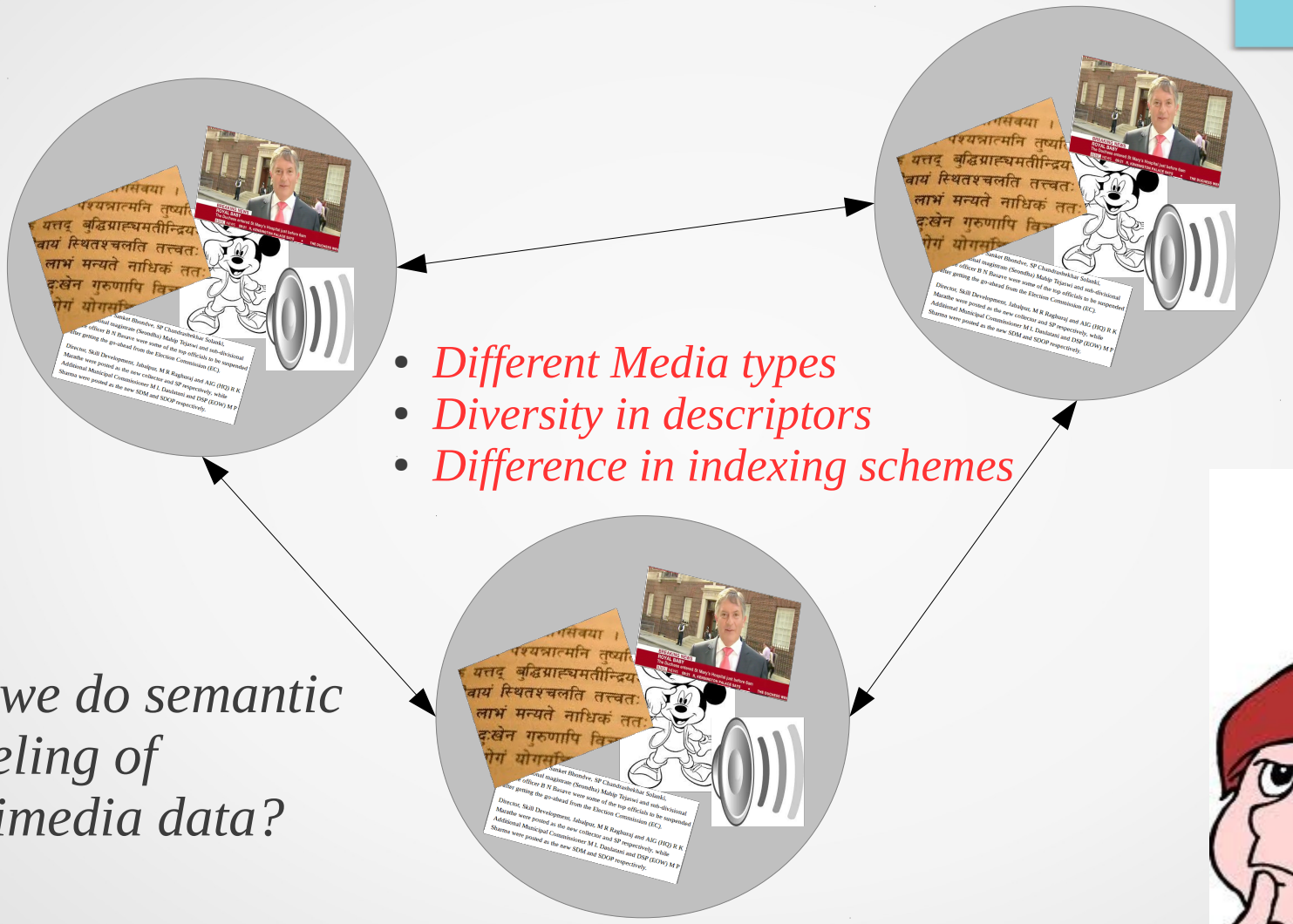
Part I

- Introduction
- Semantic Web and Ontology
- Multimedia Content Processing
- **Ontology for Multimedia Data Interpretation**

Part II

- Multimedia Web Ontology Language
- Application Examples
- Distributed Multimedia Applications
- Conclusion

Multimedia Data Integration



- *Different Media types*
- *Diversity in descriptors*
- *Difference in indexing schemes*

Can we do semantic modeling of multimedia data?

Working with the annotations

- Multimedia data is often associated with annotation
  - *Structured metadata, User tags, HTML <ALT> tag, surrounding text, ...*
- We can use ontology to interpret them?

- A set of collaborating museums *CIDOC: Early 2000's*
  - Well-curated media archives
  - Controlled metadata associated with media artifacts
- OWL-based domain ontology for information integration

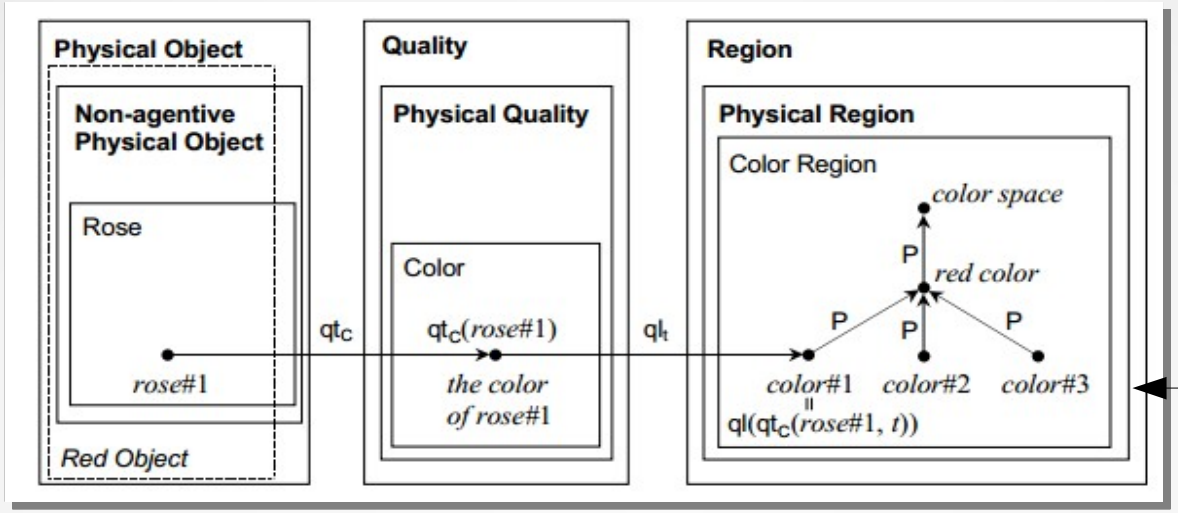
- *Unfortunately, it does not work with any arbitrary media collection*

Crowd-sourced data and knowledge *(2008 onwards)*

- Semantics extracted
  - From Crowd-sourced tags
  - With Crowd-sourced knowledge (Wikipedia)
  - *A new line of research*

- But ...
  - *Estimated 70% of social media contents are without tags*
  - *Automatic tagging*

“Qualities” of concepts



Source: Gangemi (2002)

Different shades of red

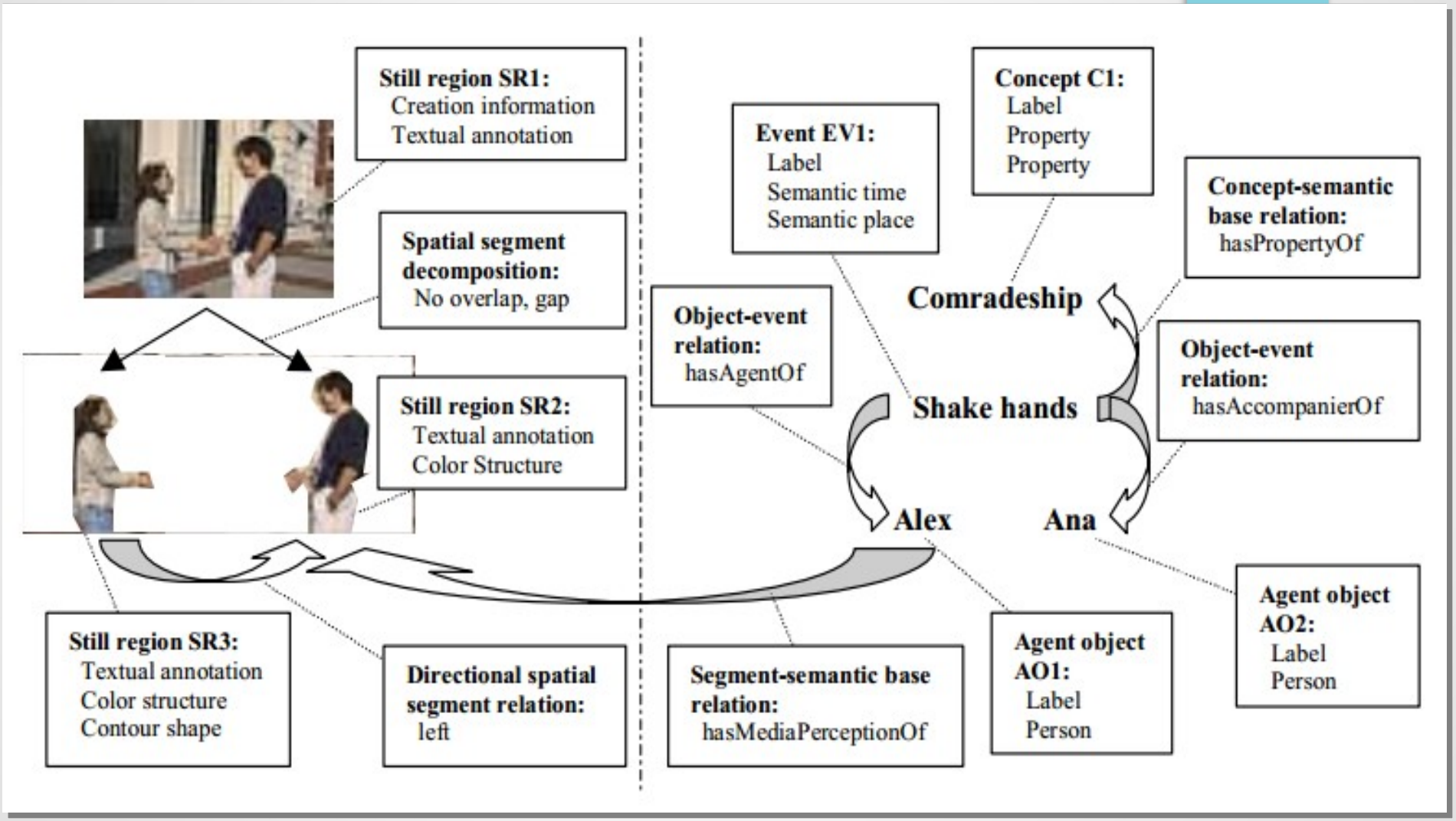
- “Qualities”: perceptible/measurable
  - Physical (color, size …)
  - Relations (Spatial and temporal)
- Relation between and quality regions (qualia)
  - “Red” is *opposite to* “green”
  - “Red” is *close to* “brown”

Multimedia Content Description Scheme

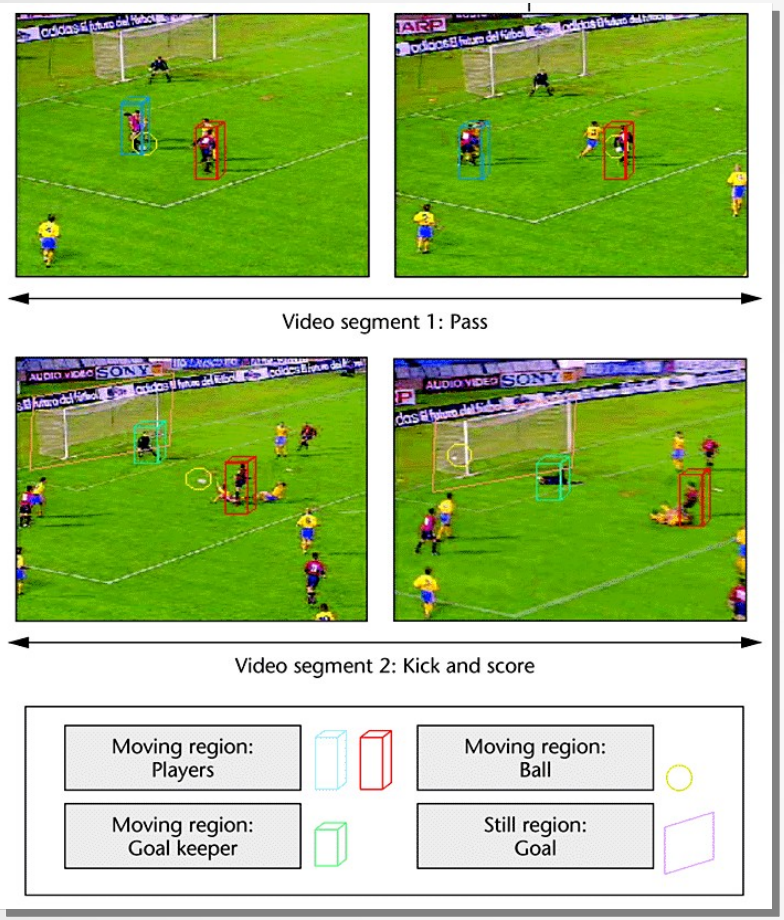
ISO Standard: MPEG-7: Early 2000's

- Flexible language to describe multimedia contents
  - Representations (tools) for common audio and visual features
    - *Color, texture, shape, frequency spectrum, etc.*
  - Scene description
    - *Structural and semantic description*
  - Extensible
    - *Possible to define new descriptors*

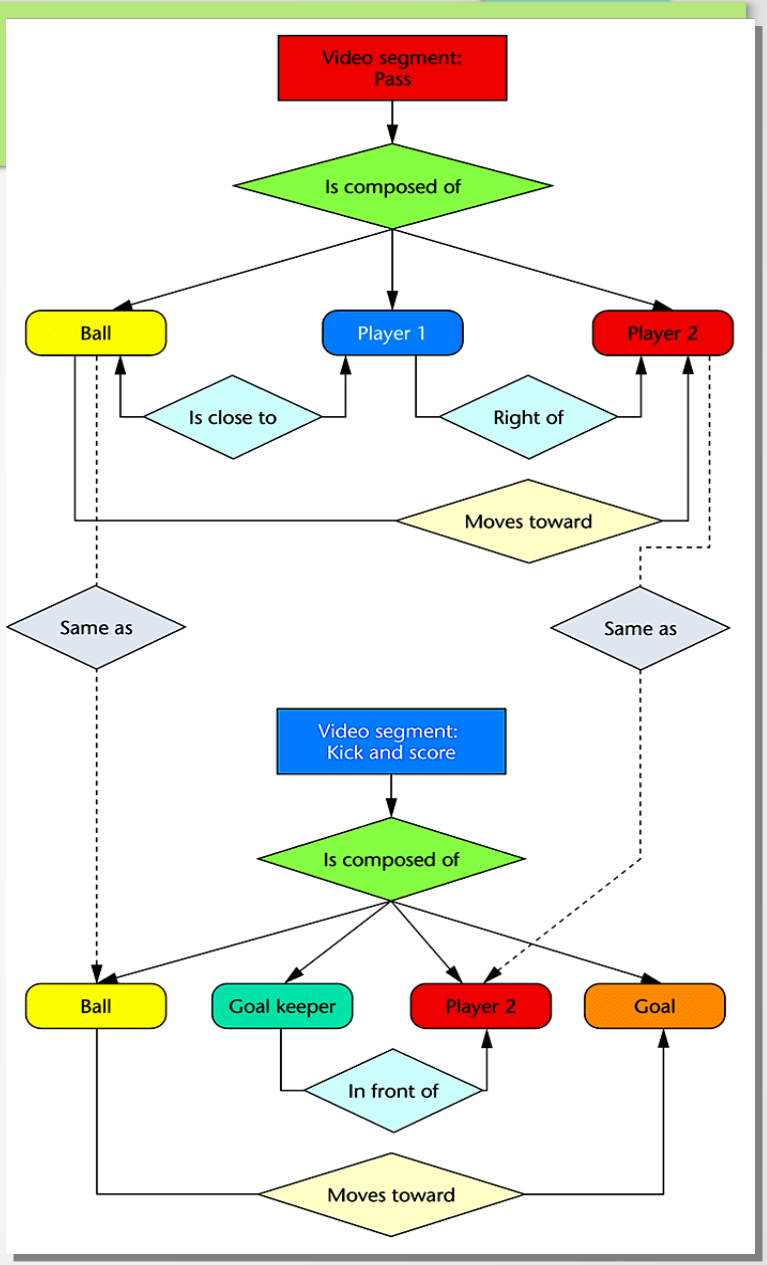
Description of still image



Video segments



Video segments and regions



Segment-Relationship Graph

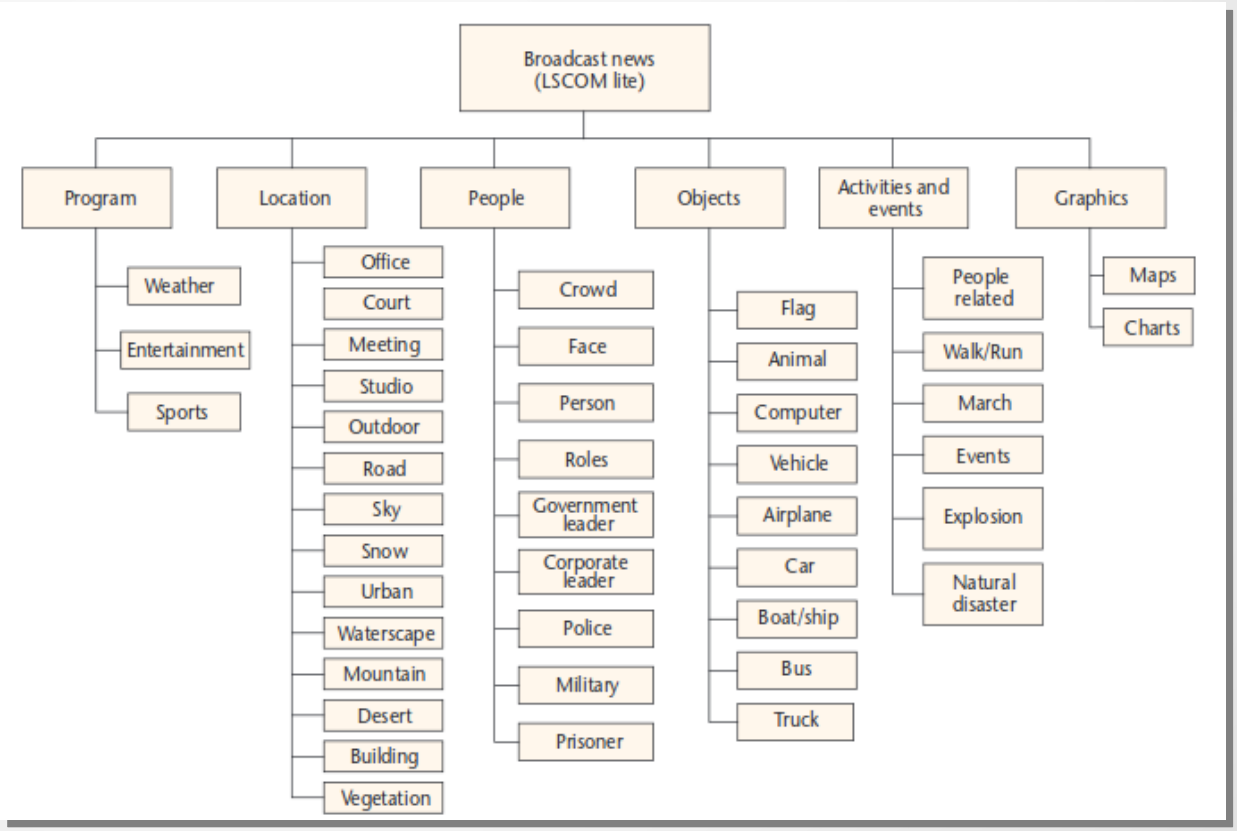
Comments on MPEG-7

- *Accomplishes syntactic interoperability for multimedia*
- Describes multimedia document content
  - XML based schema
  - Lots of flexibility (same scene can be described in many different ways)
  - No semantics, no support for reasoning
- Quite a few MM Information system built with MPEG-7
  - Template matching (query by example paradigm)

Ontology for multimedia “concepts”

IBM + CMU
Mid 2000's

- Controlled vocabulary for MPEG-7 semantic description
  - *Utility*
  - *Coverage*
  - *Feasibility*
  - *Observability*



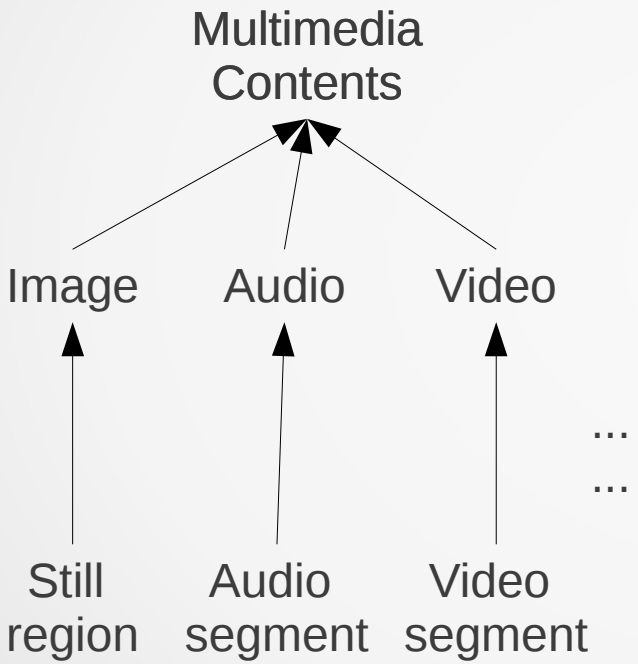
Source: Naphade (2006)

MPEG-7 Ontologies

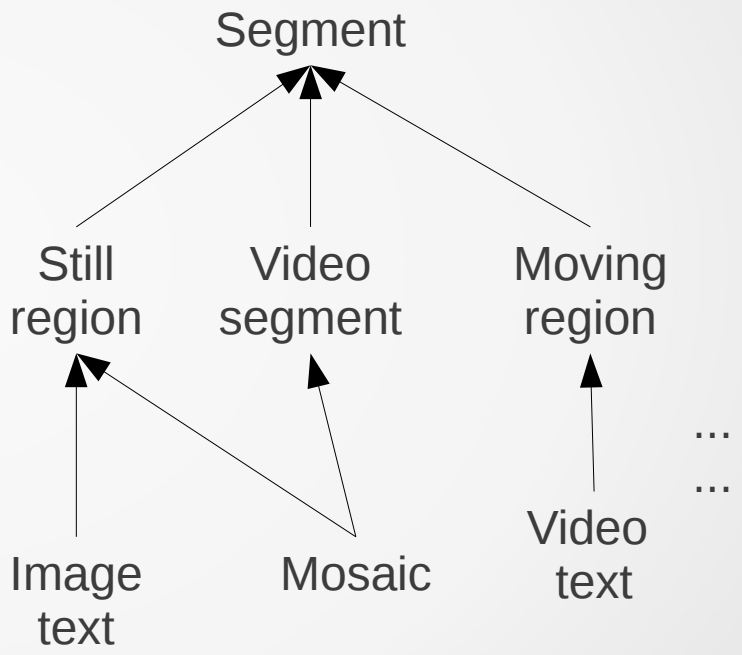
Early-Mid 2000's

- To provide semantic rigor to MPEG-7 descriptors
- Several research projects
  - Harmony
  - AceMedia
  - DS-MIRF
  - COMM
  - Boemie
  - ...
- Converts MPEG-7 constructs to RDF / OWL constructs
- Different coverage to MPEG-7 parts

MPEG-7 Ontology: Class hierarchies

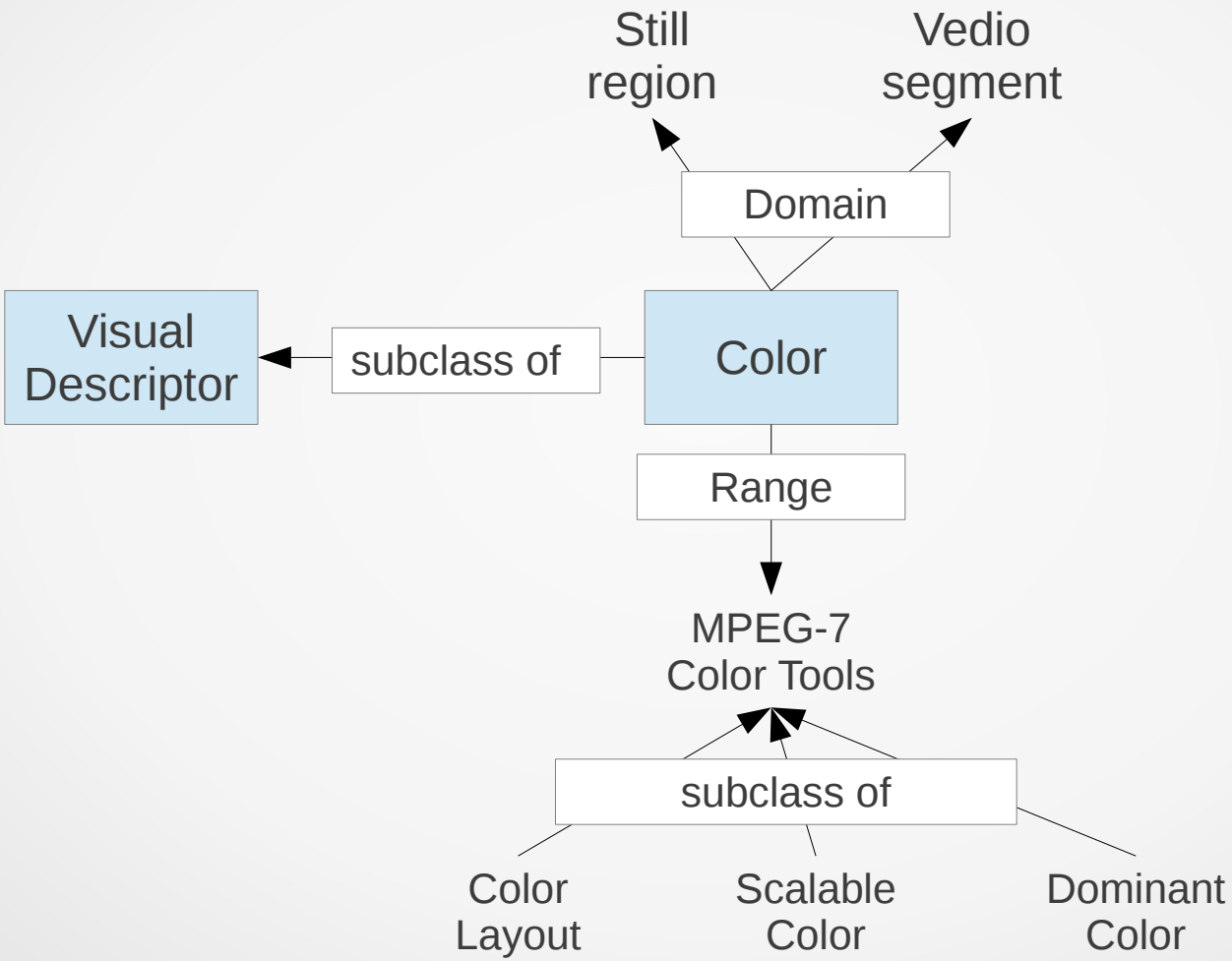


Top level content entities



Segment classes

MPEG-7 Ontology: Media Properties



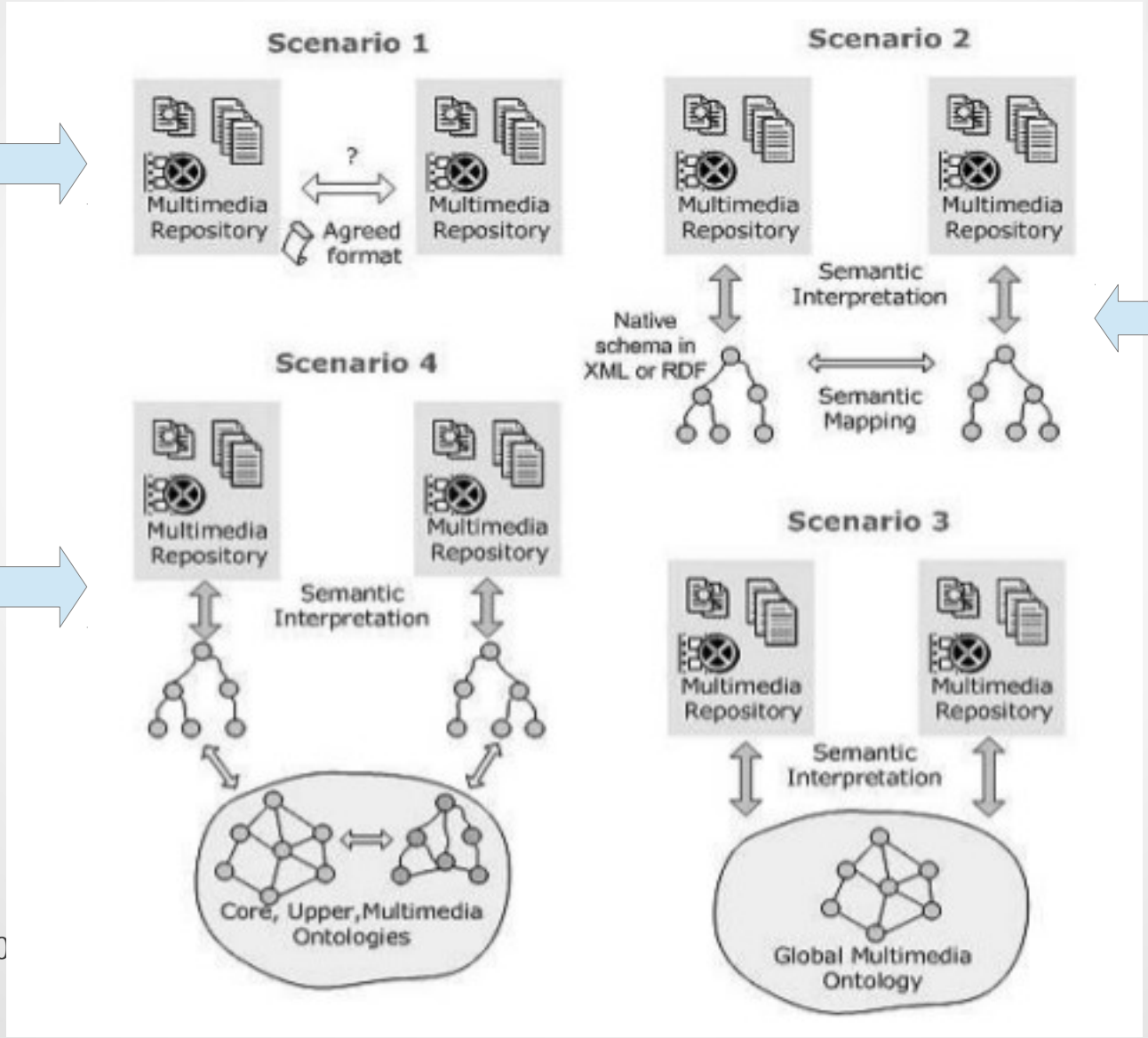
MPEG-7 Ontologies ... contd.

- Creates semantic description of multimedia contents in collections
  - **Excludes semantic descriptors**
    - Integrates with domain ontology
      - Usually with a core ontology
    - *Examples: Harmony, AceMedia, COMM*
  - **Includes semantic descriptors**
    - Results in independent semantic descriptions of repositories
    - Needs common understanding of domain
    - *Example: DS-MIRF*

Interoperability models

MPEG-7

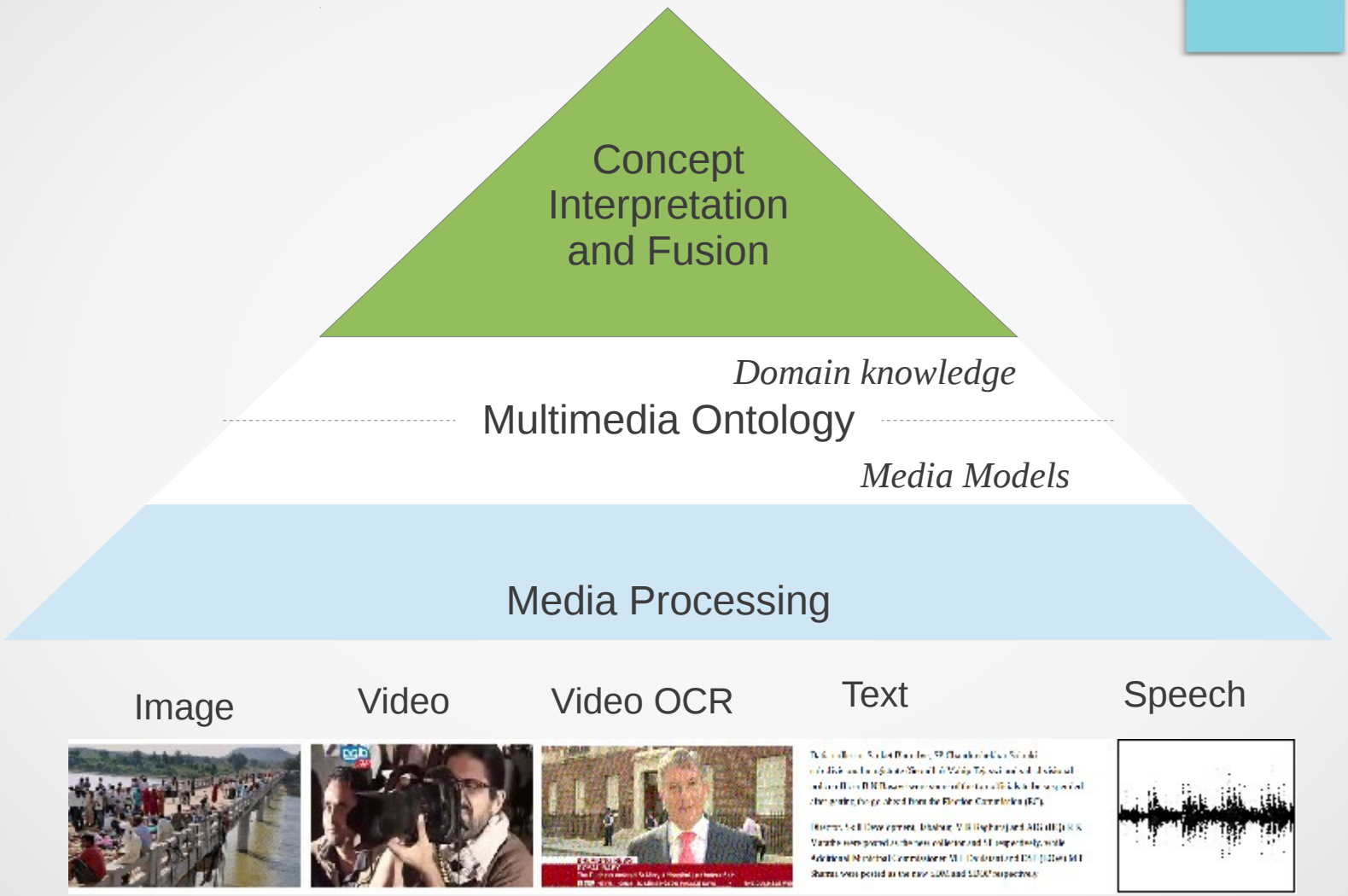
MPEG-7 Ontology
- *Harmony*
- *AceMedia*
- *COMM*



MPEG-7 Ontology
- *DS-MIRF*

Source: Dasiopoulou (2010)

Architecture for semantic integration



Comments on MPEG-7 ontologies

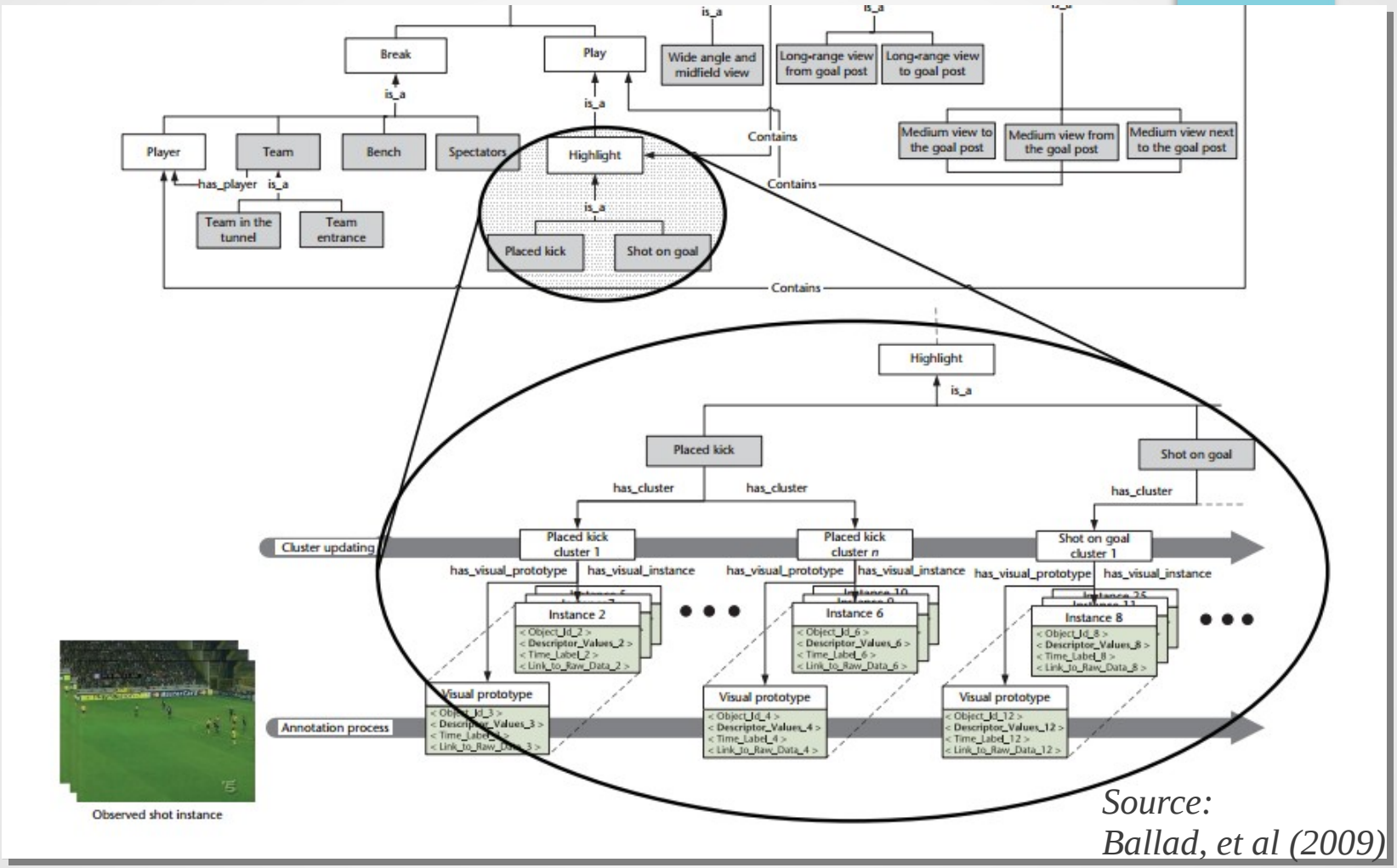
- Content model for documents / collections
  - Can correlate diverse media forms
- Specific to multimedia instances
  - Not a generic collection independent ontology
- Media model and domain model form separate layers
  - Media interpretation does not benefit from domain knowledge

Pictorially enhanced ontology

Univ Fierenze: Mid-Late 2000's

- Visual templates (examples) are associated with media events (concepts)
  - Each template represents a distinct modality of manifestation
- New instances are classified based on feature similarity with prototypes
  - Automatic event detection and annotation
- Domain ontology relates such events

Pictorially enhanced ontology ... contd



Source: Ballad, et al (2009)

End of Part I