# A Semantic Multimedia Web: Create, Annotate, Present and Share your Media

Raphaël Troncy, Lynda Hardman
<Raphael.Troncy|Lynda.Hardman@cwi.nl>

CWI, Semantic Media Interfaces
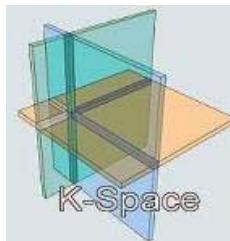
Learning Objectives

- Understand multimedia applications workflow
  - Take the canonical processes of media production model
- Explore various multimedia metadata formats
  - Be aware of the advantages and limitations of various models
  - Know the interoperability issues and understand COMM, a Core Ontology for Multimedia
- Discuss exploratory interfaces based on rich multimedia metadata semantics
  - Know how to link and expose your data on the web
  - See various multimedia presentation interfaces

Agenda

1. Understanding Multimedia Applications Workflow

    – CeWe Color Photo Book creation application

    – Vox Populi argumentation-based video sequences generation

    – *Canonical Processes of Media Production*

2. Semantic Annotation of Multimedia Content

    – Multimedia metadata formats: use cases and requirements

    – Multimedia metadata interoperability issues

    – MPEG-7 based ontologies

    – *COMM: A Core Ontology for MultiMedia*

3. Semantic Search and Presentation of Multimedia Content

    – Link your data!

    – *Searching and Browsing Multimedia Semantic Datasets with Cliopatria*

Understanding Multimedia Applications Workflow

- Identify and define a number of canonical processes of media production

- Community effort

  - 2005: [Dagstuhl seminar]

  - 2005: ACM MM Workshop on *Multimedia for Human Communication*

  - 2008: Multimedia Systems Journal Special Issue (core model and companion system papers)
    editors: Frank Nack, Zeljko Obrenovic and Lynda Hardman

Overview of Canonical Processes
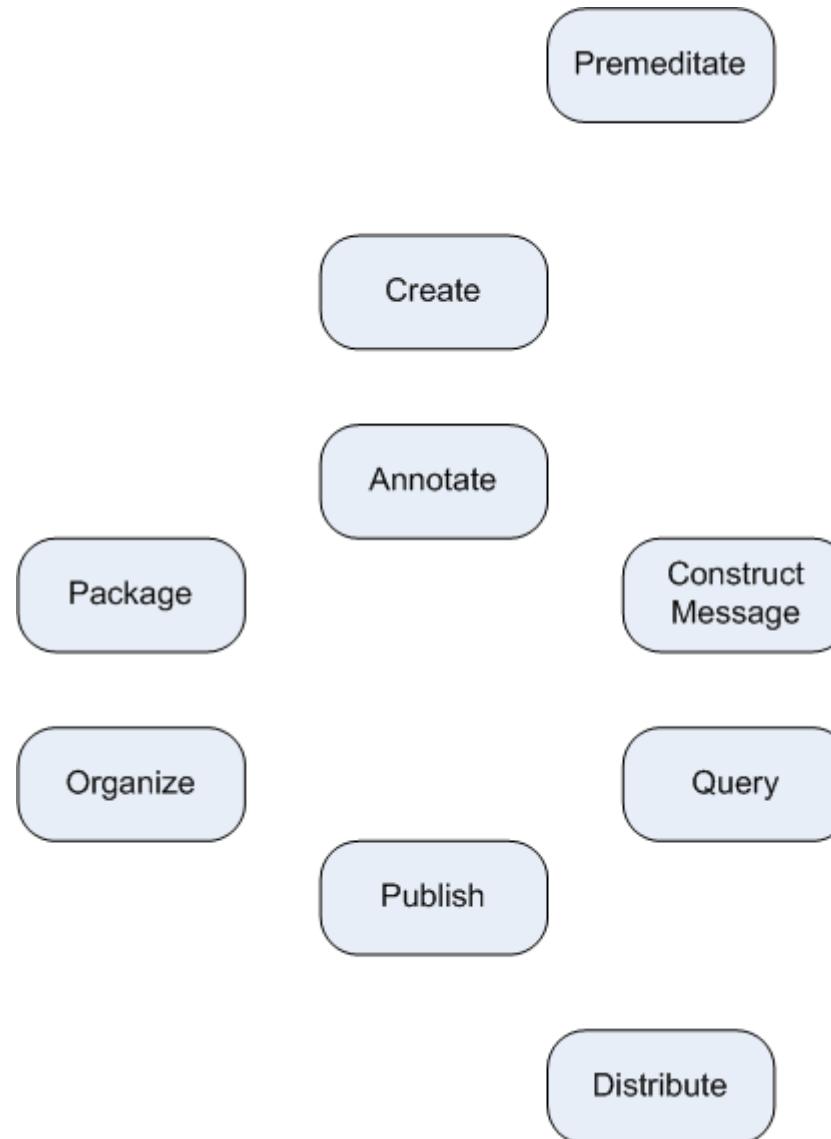
Example 1: CeWe Color PhotoBook

- Application for authoring digital photo books
- Automatic selection, sorting and ordering of photos
  - Context analysis methods: timestamp, annotation, etc.
  - Content analysis methods: color histograms, edge detection, etc.
- Customized layout and background
- Print by the European leader photo finisher company

http://www.cewe-photobook.com

CeWe Color PhotoBook Processes

- My winter ski holidays with my friends
-
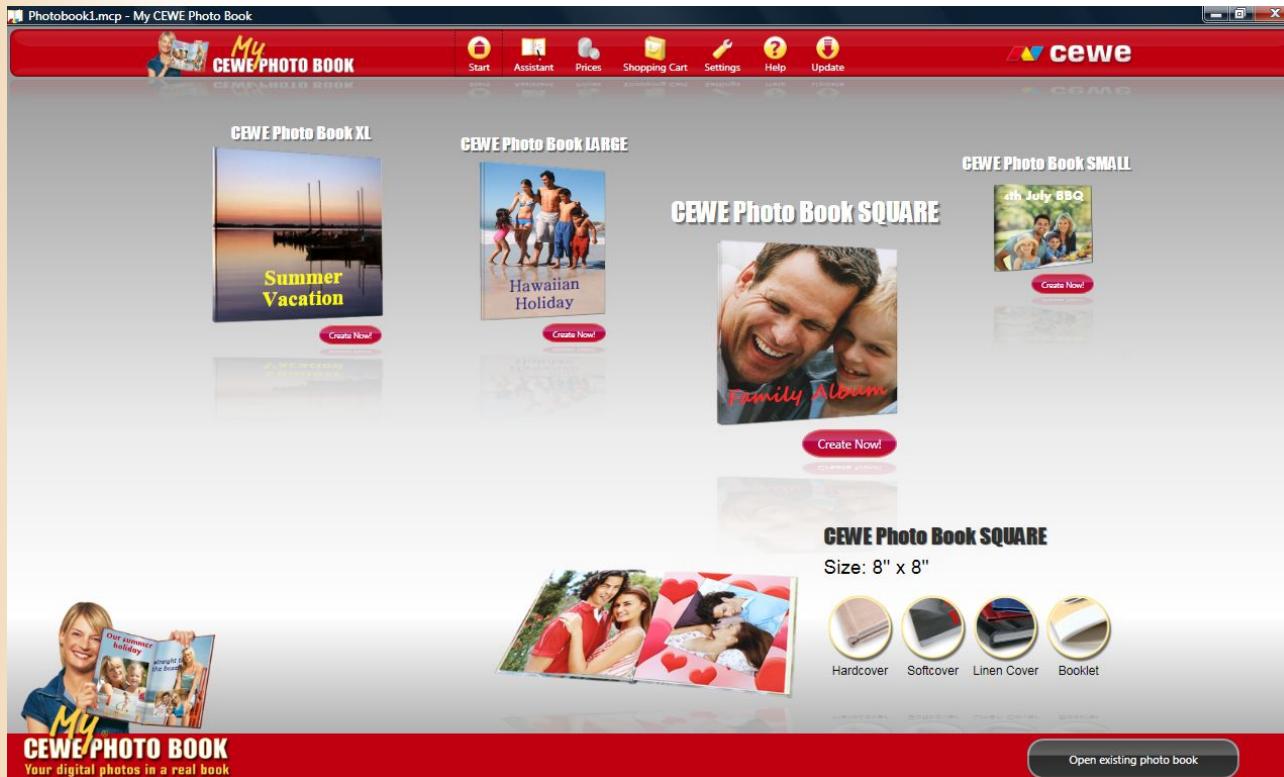
Premeditate

Construct Message

# CeWe Color PhotoBook Processes
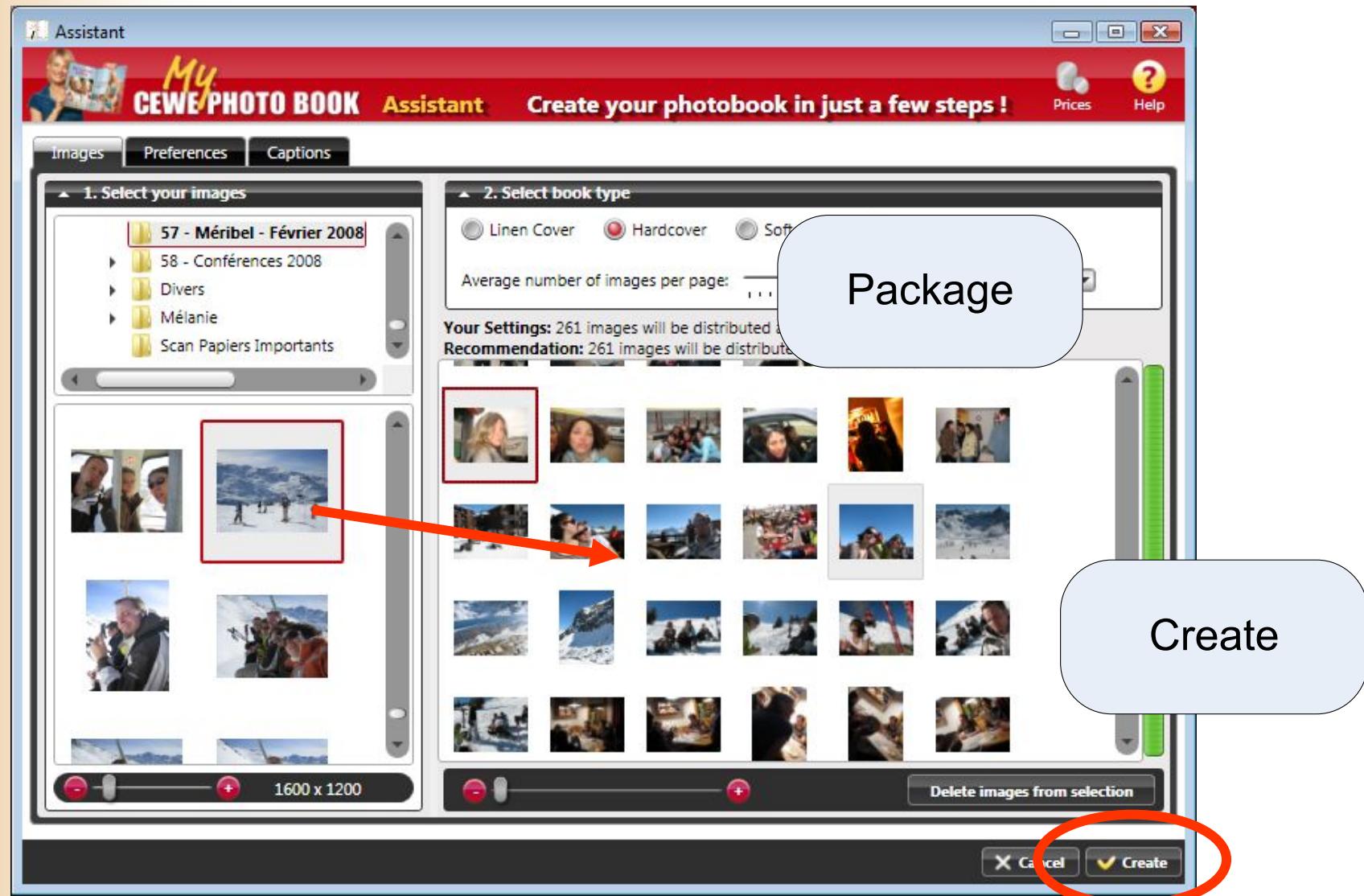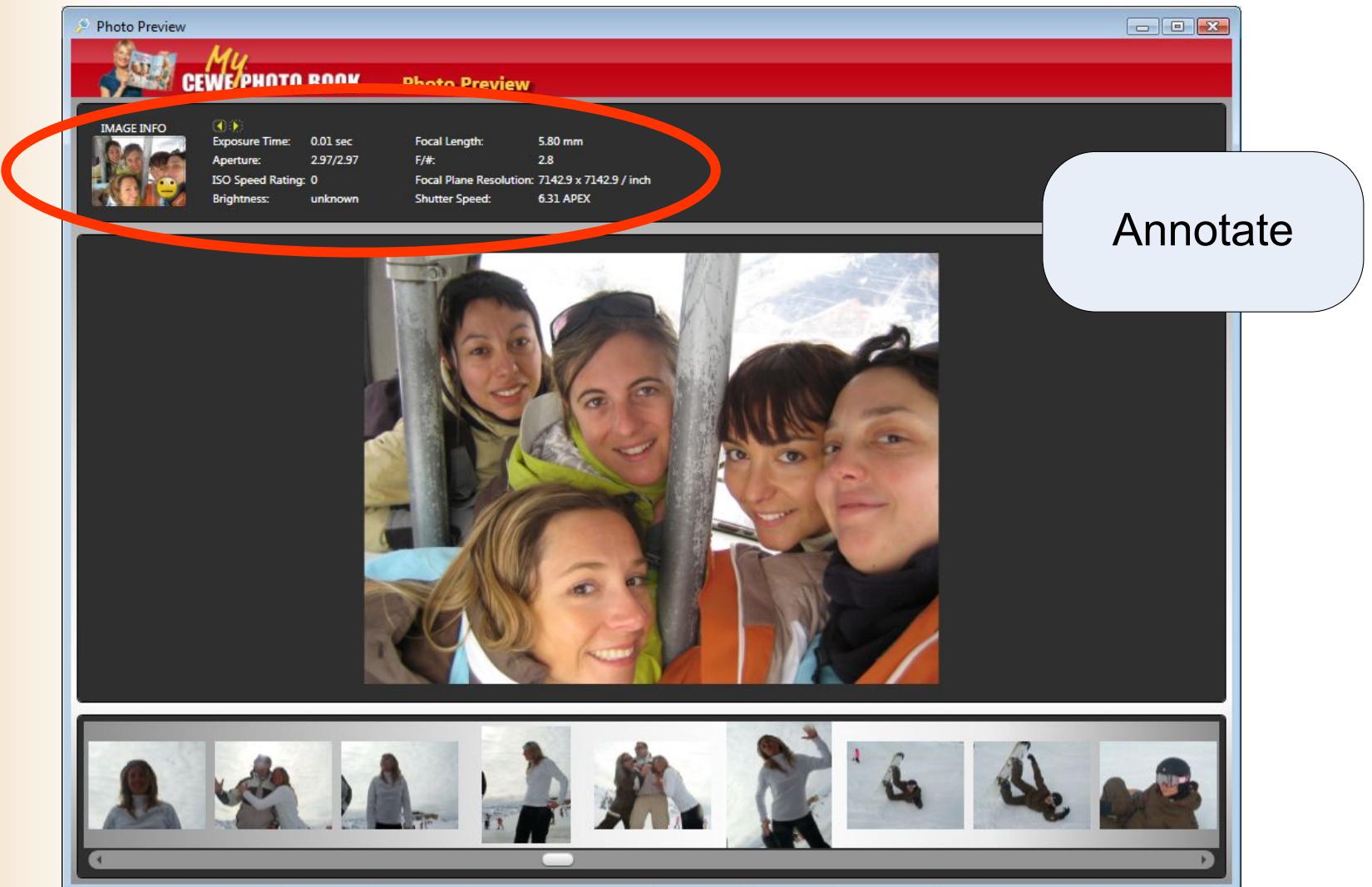
CeWe Color PhotoBook Processes

-

CeWe Color PhotoBook Processes

CeWe Color PhotoBook Processes

-

Publish

Distribute
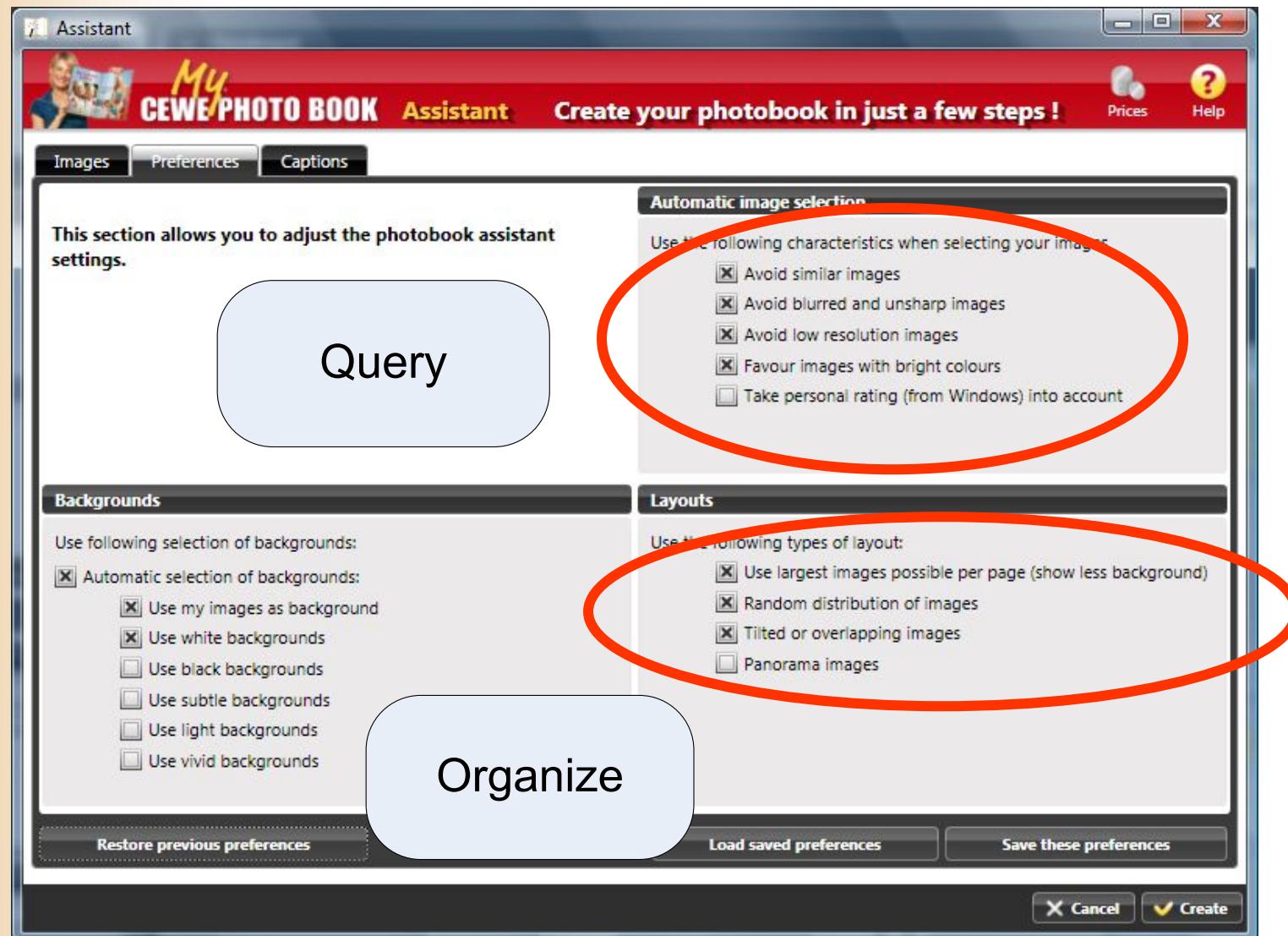
CeWe Color PhotoBook Processes
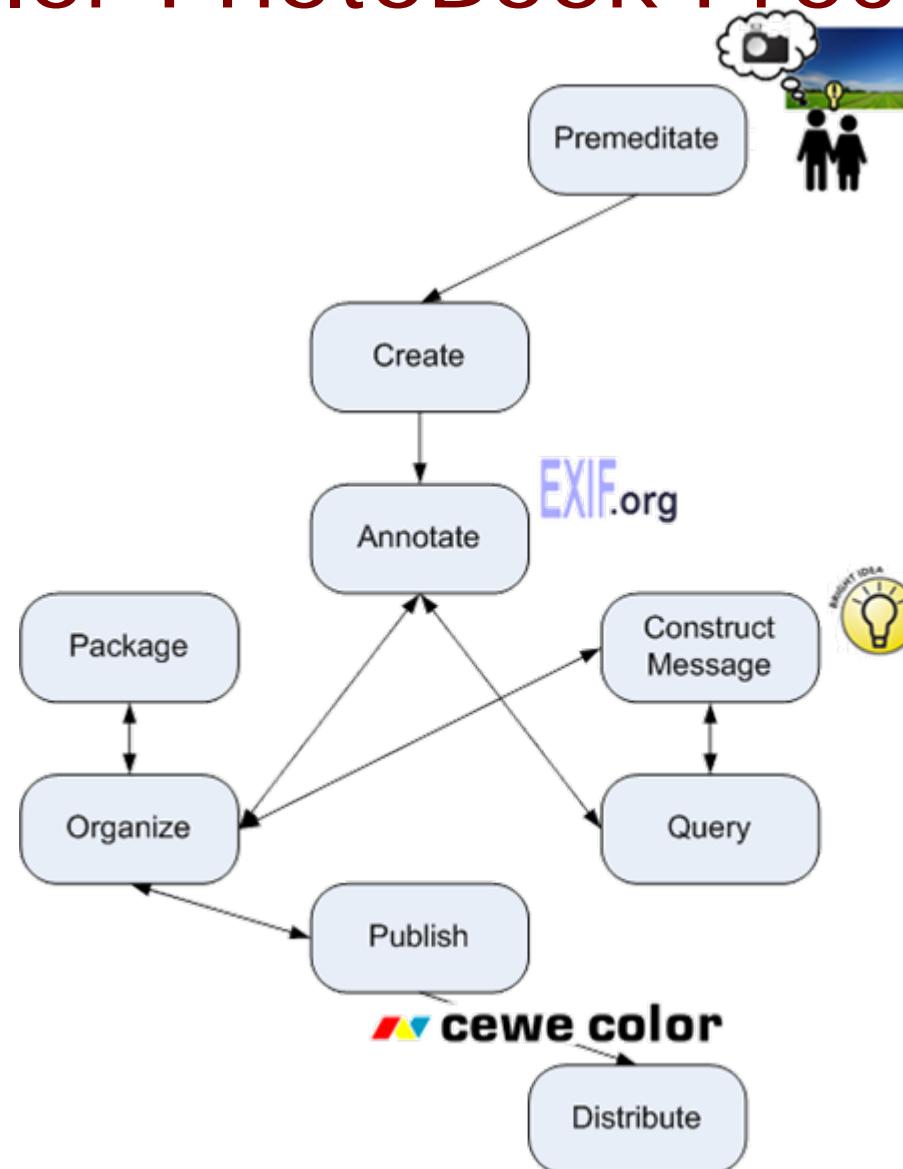
Example 2: Vox Populi Video Sequences Generation

Stefano Bocconi, Frank Nack

- **Interview with America**
  video footage with interviews and background material about the opinion of American people after 9-11
  http://www.interviewwithamerica.com

- Example question:
  What do you think of the war in Afghanistan?

"I am never a fan of military action, in the big picture I don't think it is ever a good thing, but I think there are circumstances in which I certainly can't think of a more effective way to counter this sort of thing..."

Vox Populi Premeditate Process

- Analogous to the pre-production process in the film industry

  - *Static* versus *dynamic* video artifact

  Premeditate

- Output

  - Script, planning of the videos to be captured

  - Questions to the interviewee prepared

  - Profiles of the people interviewed: education, age, gender, race

  - Locations where the interviews take place

Vox Populi Annotations

- Contextual
  - Interviewee (social), locations
- Descriptive
  - Question asked and transcription of the answers
  - Filmic continuity, examples:
    - gaze direction of speaker (left, centre, right)
    - framing (close-up, medium shot, long shot)
- Rhetorical
  - Rhetorical Statement
  - Argumentation model: Toulmin model

Vox Populi Statement Annotations

- Statement formally annotated:
  - *<subject> <modifier> <predicate>*
  - E.g. "**war best solution**"

- A thesaurus containing:
  - Terms on the topics discussed (155)
  - Relations between terms: *similar* (72), *opposite* (108), *generalization* (10), *specialization* (10)
  - E.g. **war** *opposite* **diplomacy**

Toulmin Model
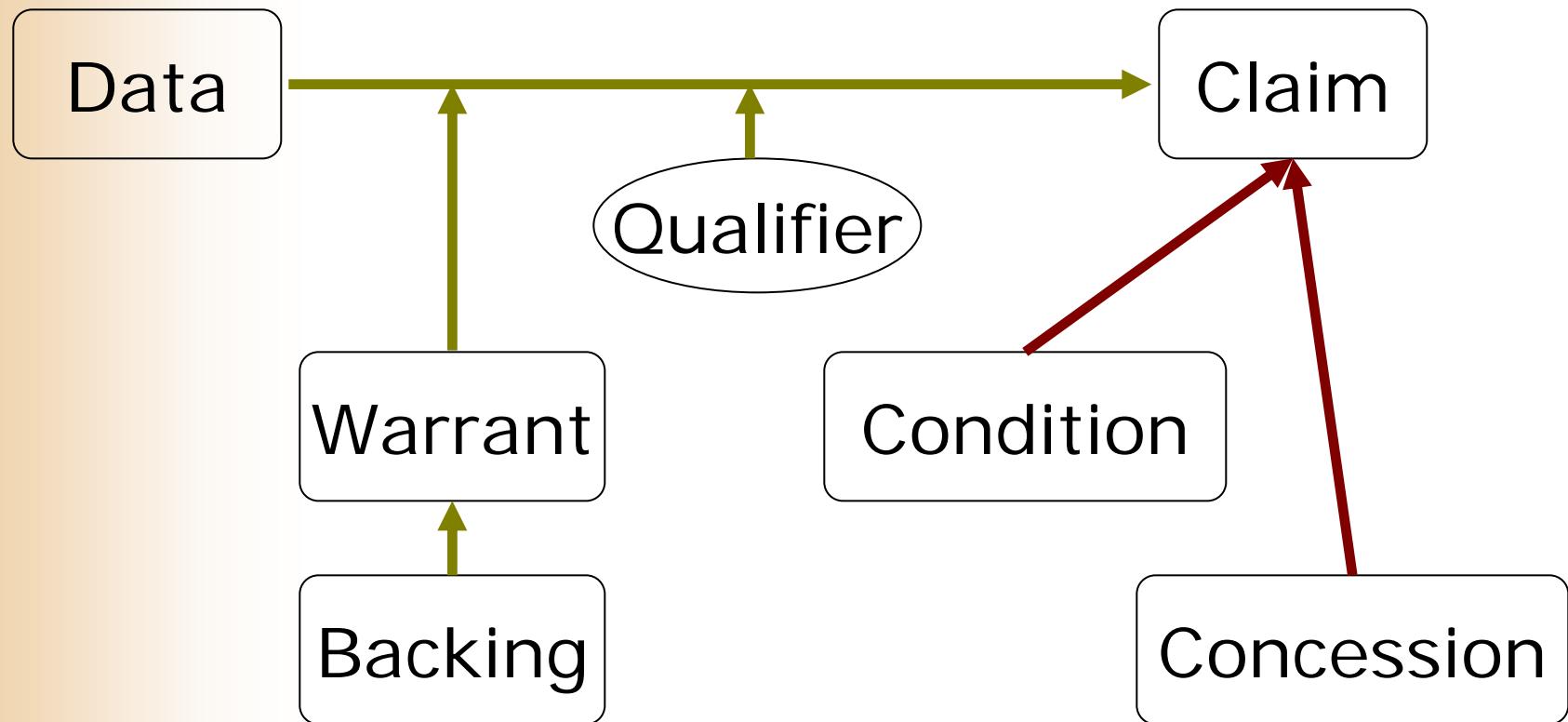
57 Claims, 16 Data, 4 Concessions, 3 Warrants, 1 Condition

Vox Populi Query Interface

Construct Message

| Question	Interviewee	Opinion
Why did they do what they did?	Cameroun Parking Guard at Stamford	War in Afghanistan - Pro
What do you think of the casualties among civilians?	Lawyer in Harward	
(What do you think of the Afghanistan war?)		
What are the consequences of the war?		
What are the roots of the problem?		
What do you think about the Anthrax?		

Position

First Character

Age	Education	Employment	GeoLocation	Race	Religion	Sex
Middleage	HighEducated	HighIncomeJob	NotUSA	AmericanIndian	Atheist	Female
Old	LowEducated	LowIncomeJob	USA	Asian	Christian	Male
Teenager	MediumEducated	MiddleIncomeJob		Black	Muslim	
Young		Retired		Hispanic		
		Student		White		

Second Character

Age	Education	Employment	GeoLocation	Race	Religion	Sex
Middleage	HighEducated	HighIncomeJob	NotUSA	AmericanIndian	Atheist	Female
Old	LowEducated	LowIncomeJob	USA	Asian	Christian	Male
Teenager	MediumEducated	MiddleIncomeJob		Black	Muslim	
Young		Retired		Hispanic		
		Student		White		

Strategy	Bandwidth	Intercut	Caption
○ None	○ Low Bandwidth	⊙ True	○ On (can cause problems)
⊙ Create Clash	⊙ Medium Bandwidth	○ False	⊙ Off
○ Create Support	○ High Bandwidth		
○ Vox Populi			

Query

[Done] [Reset]

Vox Populi Organize Process

- Using the thesaurus, create a graph of related statements
  - nodes are the statements (corresponding to video segments)
    "*war best solution*",
    "*diplomacy best solution*",
    "*war not solution*"
  - edges are either *support* or *contradict*

Organize

diplomacy best solution

contradict

support

war best solution

war not solution
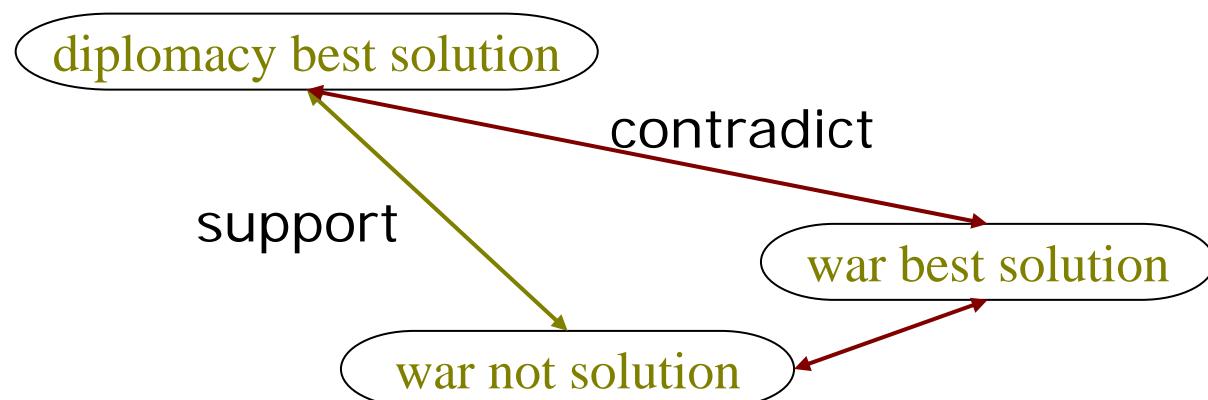
Result of Vox Populi Query

I am not a fan of military actions

**I cannot think of a more effective solution**
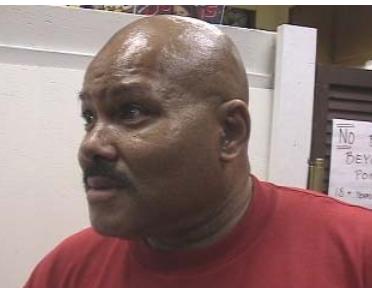
Publish

War has never solved anything

Distribute

Two billions dollar bombs on tents

Vox Populi Processes
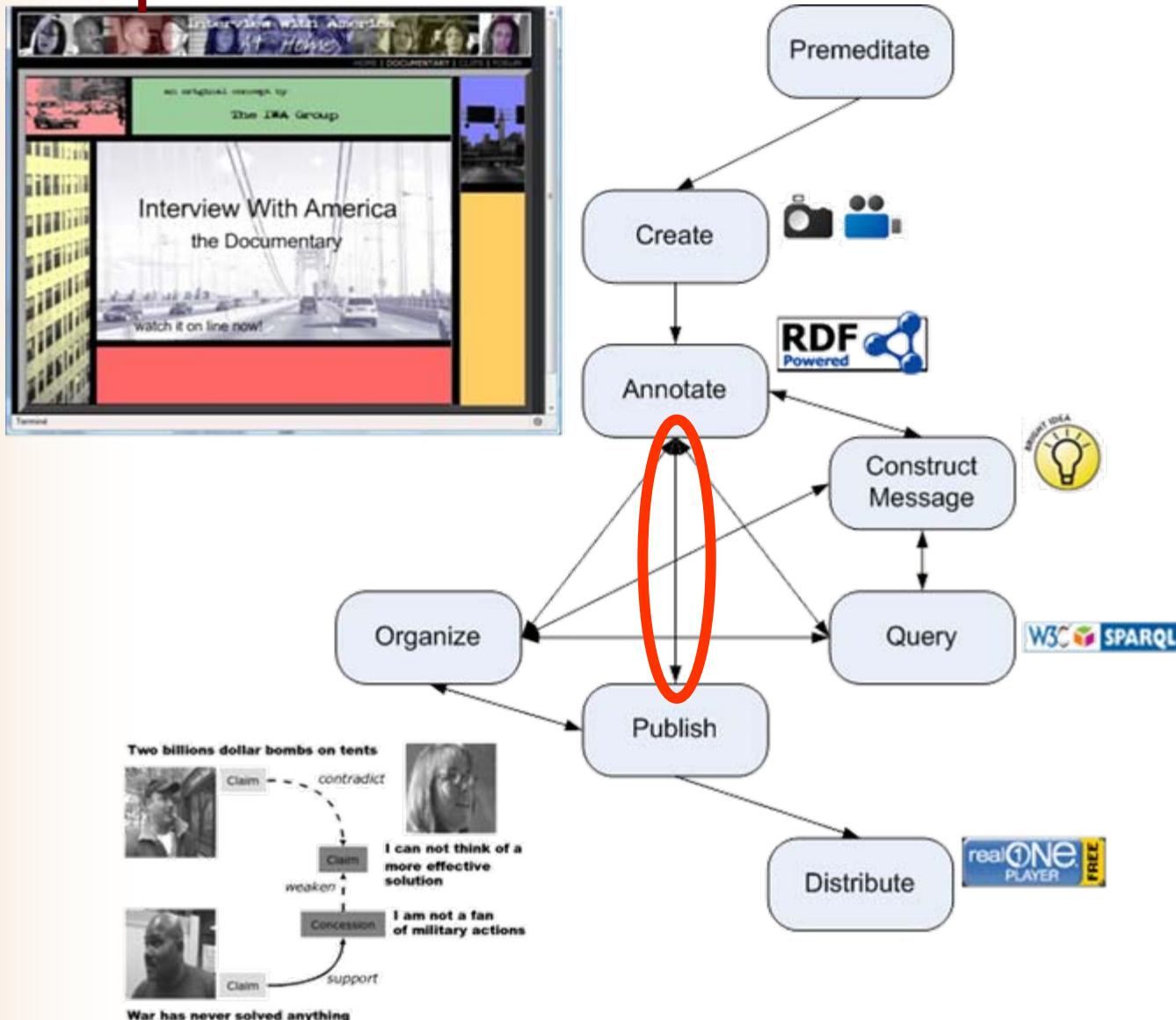
Canonical Processes 101

- Canonical: reduced to the simplest and most significant form possible without loss of generality

- Formalization of each process in UML diagrams
    - Process
    - Process artifacts
    - Process actors
    - External world artifacts

Premeditate

- Process where initial ideas about media production are established

  - *Design a photo book of my last holidays for my family*
  - *Create argument-based sequences of videos of interviews after September 11*
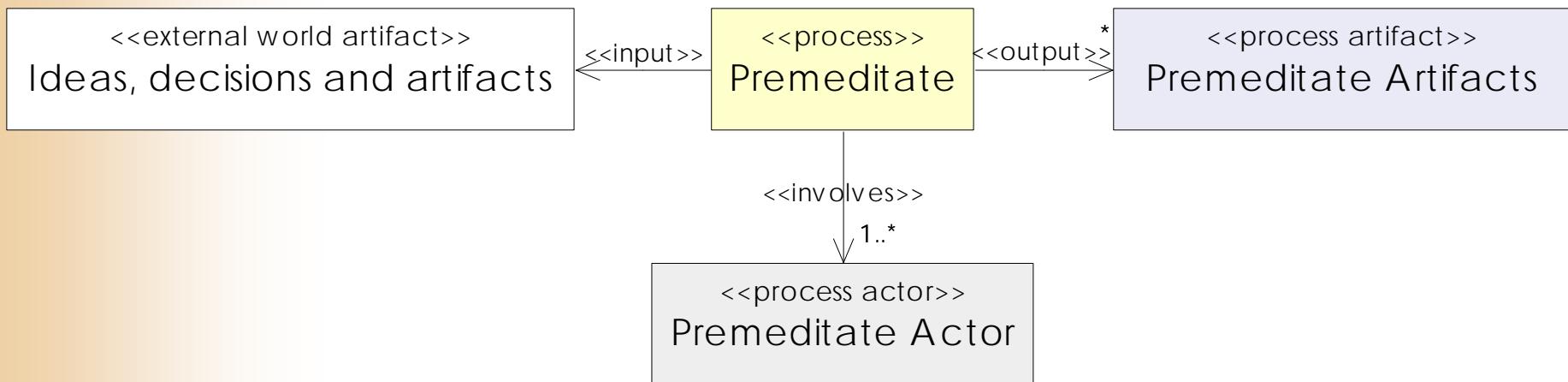
Create Media Asset

- Process where media assets are captured, generated or transformed
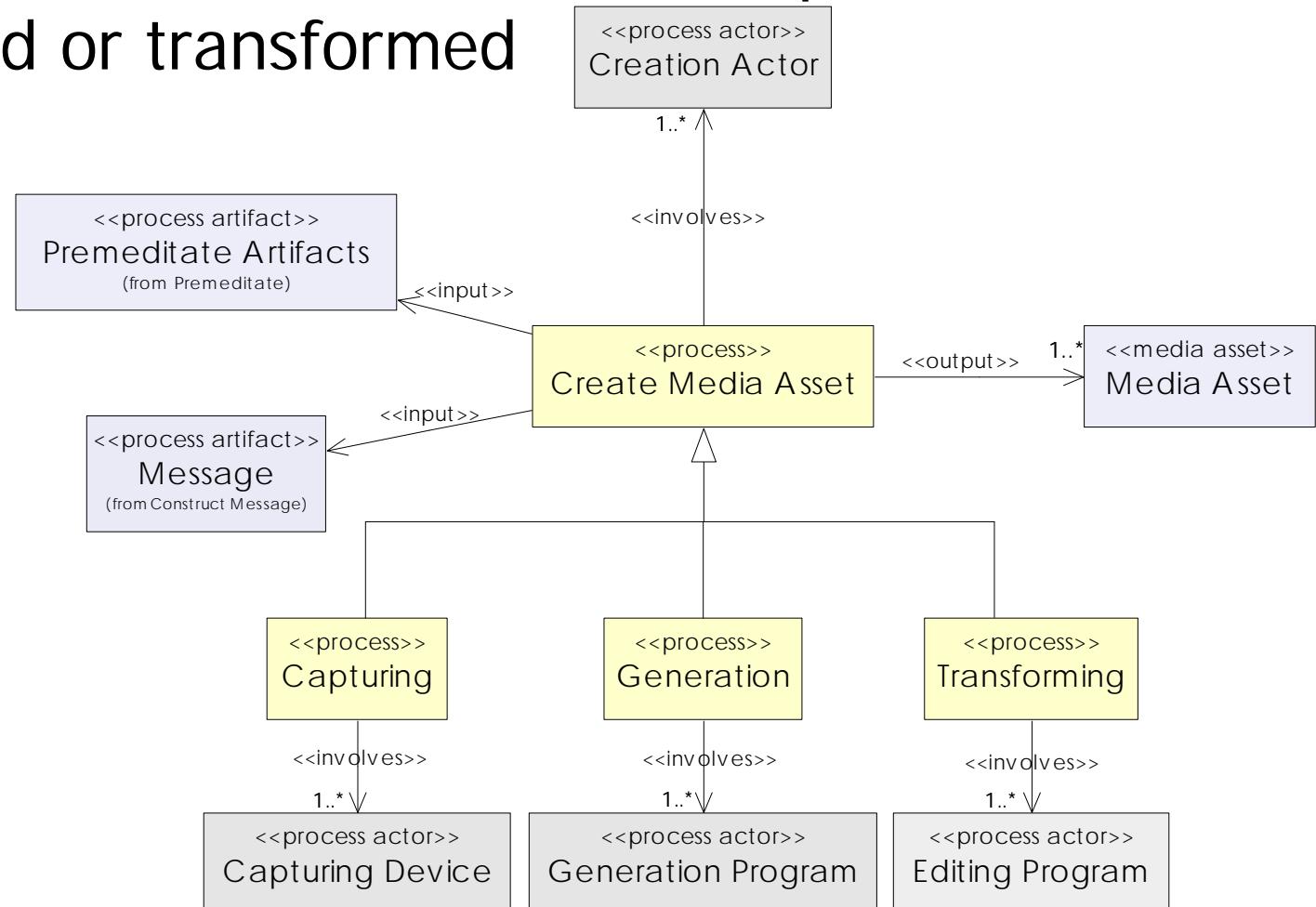
Annotate

- Process where annotation is created
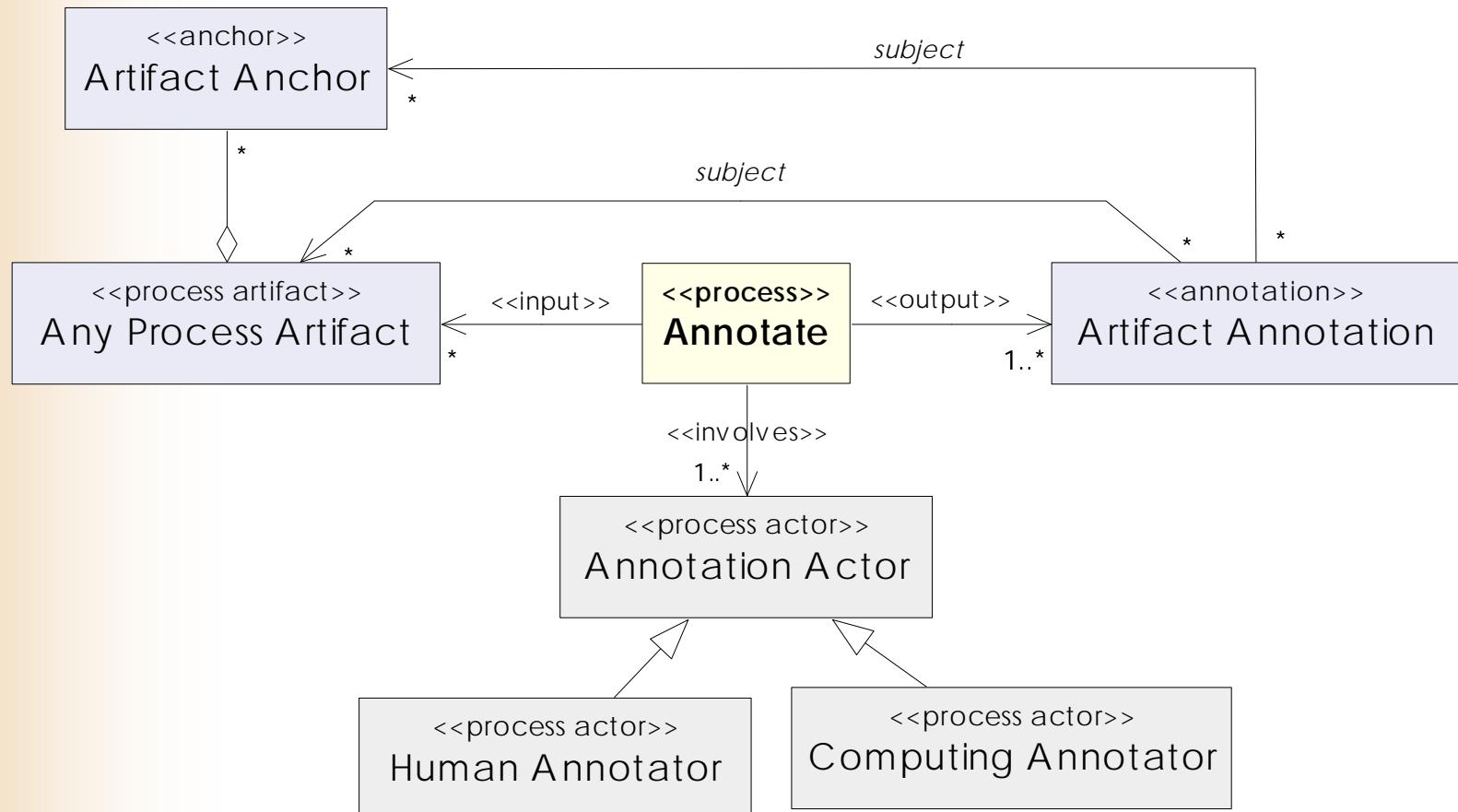
Semantic Annotate

- The annotation uses some controlled vocabularies
  - *Subject matter annotations of your photos*
  - *Rhetorical annotations in Vox Populi*

| <<process artifact>>
Any Process Artifact | <<input>> | <<process>>
Annotate | <<output>> | <<annotation>>
Artifact Annotation |

*

1..*

<<schema>>
Vocabulary

<<input>>

<<process>>
Semantic Annotate

<<output>>

<<annotation>>
Semantic Artifact Annotation

link

<<term>>
Vocabulary Term

described in terms of
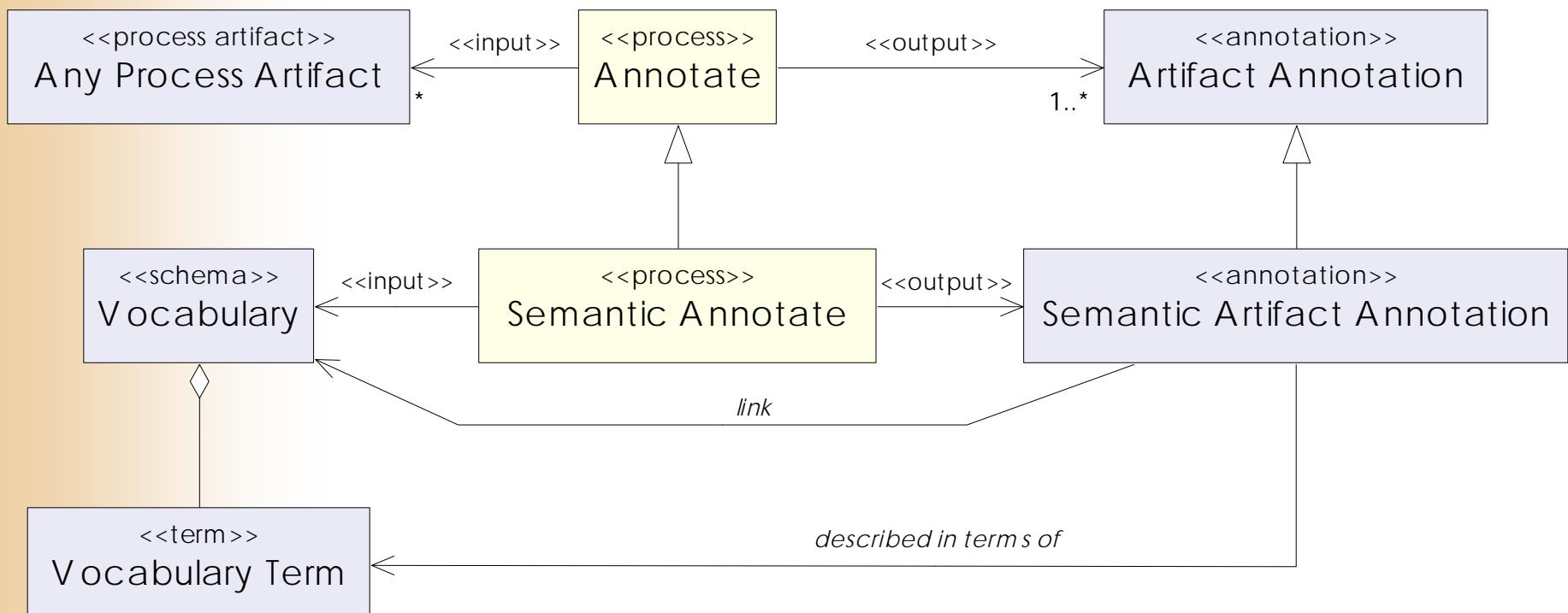
Package

- Process where process artifacts are logically and physically packed
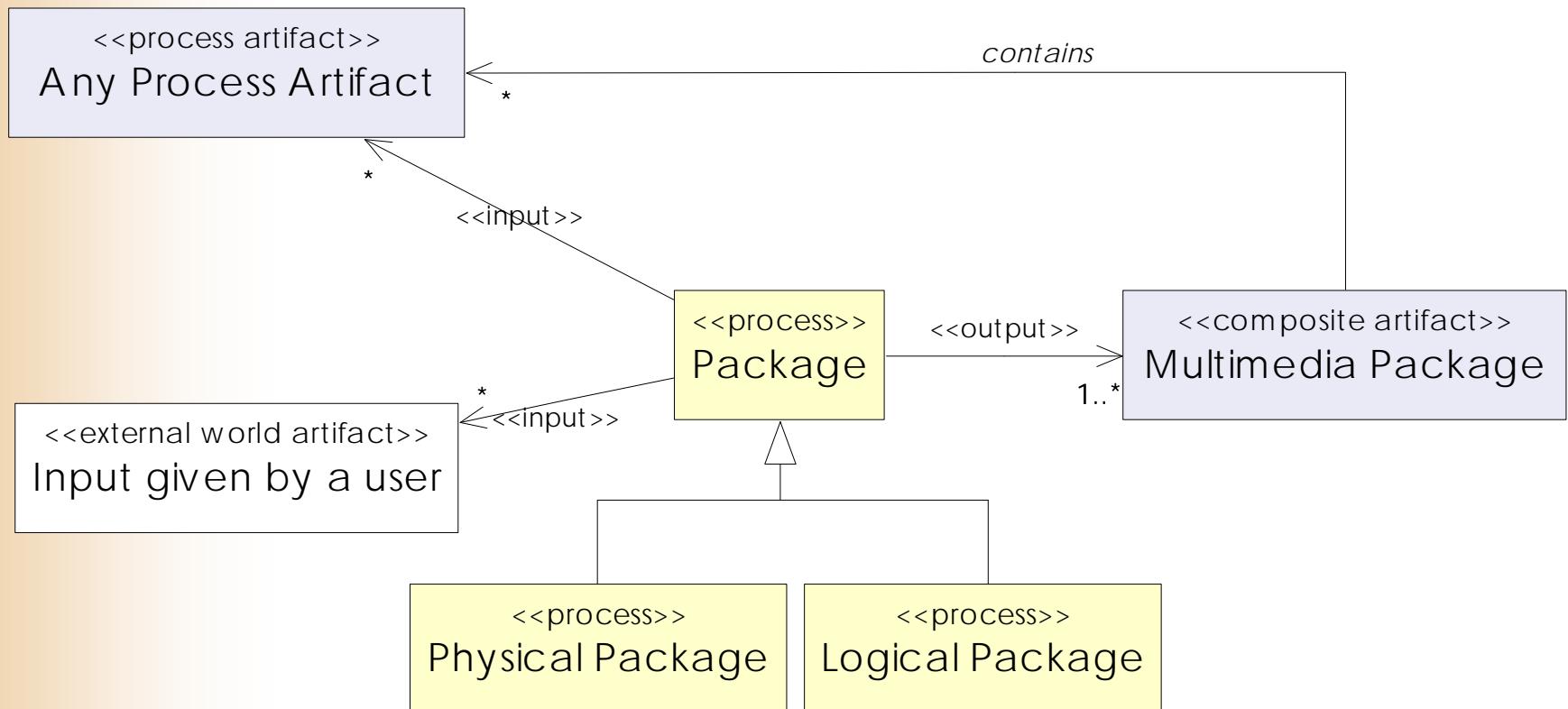
Query

- Process where a user retrieves a set of process artifacts based on a given query
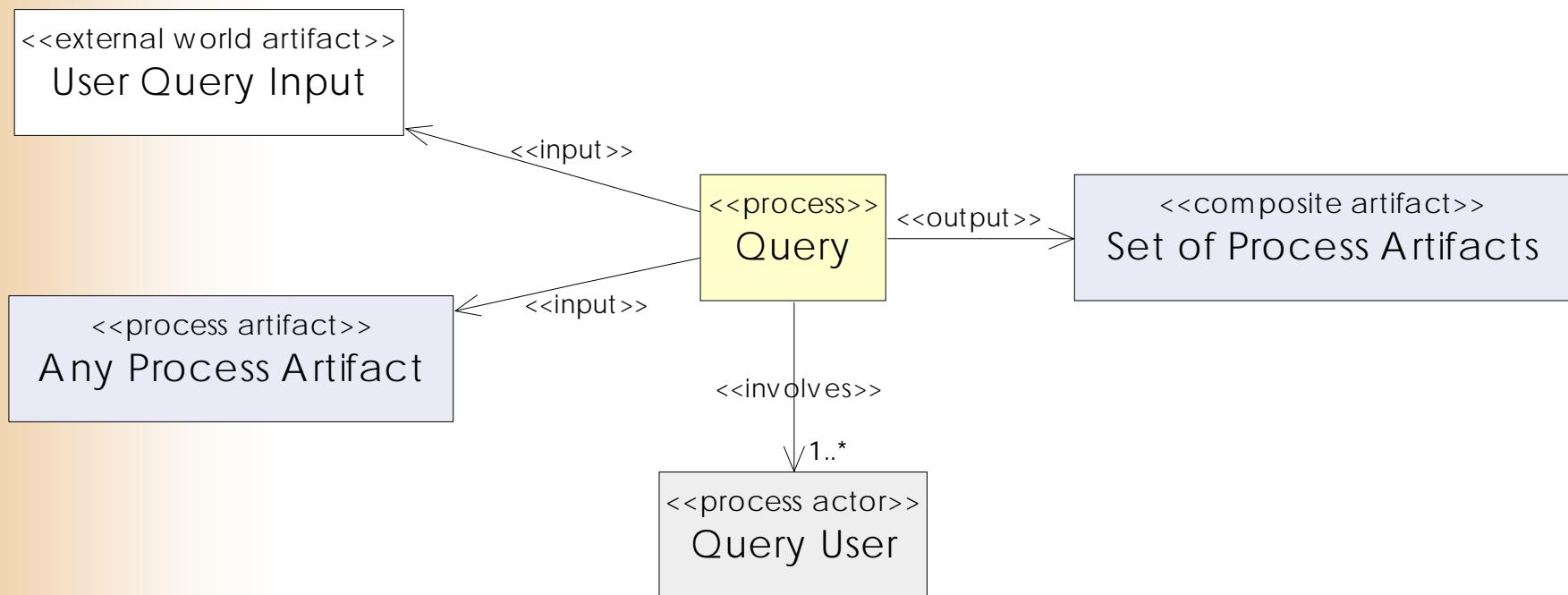
Construct Message

- Process where an author specifies the message they wish to convey
  - *Our holiday was sporty, great weather and fun*
  - *Create clash about whether war is a good thing*


```
┌────────────────────────────┐          ┌─────────────────────────┐          ┌──────────────────────────┐
│   <<external world artifact>>│ <<input>>│      <<process>>        │<<output>>│   <<process artifact>>   │
│  Ideas, decisions and artifacts│◄───────│    Construct Message    │─────────►│         Message          │
└────────────────────────────┘          └─────────────────────────┘          └──────────────────────────┘
                                                   │
                                              <<involves>>
                                              1..* ▼
                                         ┌─────────────────────────┐
                                         │     <<process actor>>   │
                                         │      Message Author     │
                                         └─────────────────────────┘
```
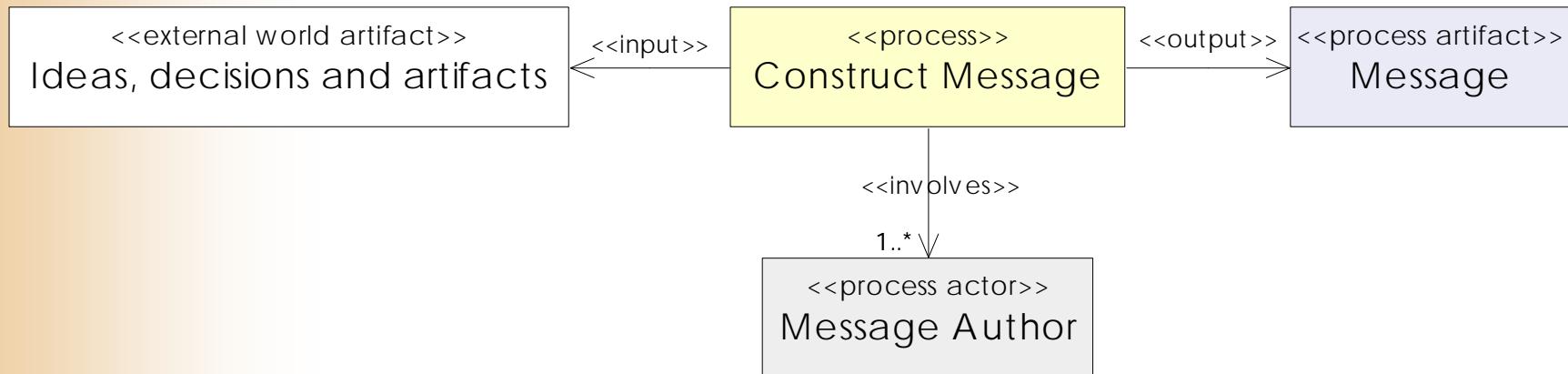
Organize

- Process where process artifacts are organized according to the message
  - *Organize a number of 2-page layouts in photobook*
  - *Use semantic graph to select related video clips to form linear presentation of parts of argument structure*

```
                <<process artifact>>
                     Message
                (from Construct Message)
                         |
                    0..1  \
                           \  <<input>>
                            \
         <<process>>              <<composite artifact>>
          Organize   --<<output>>-->  Document Structure
                    /
          <<input>>/
                  /
    <<process artifact>>        contains
    Any Process Artifact  <----------
```
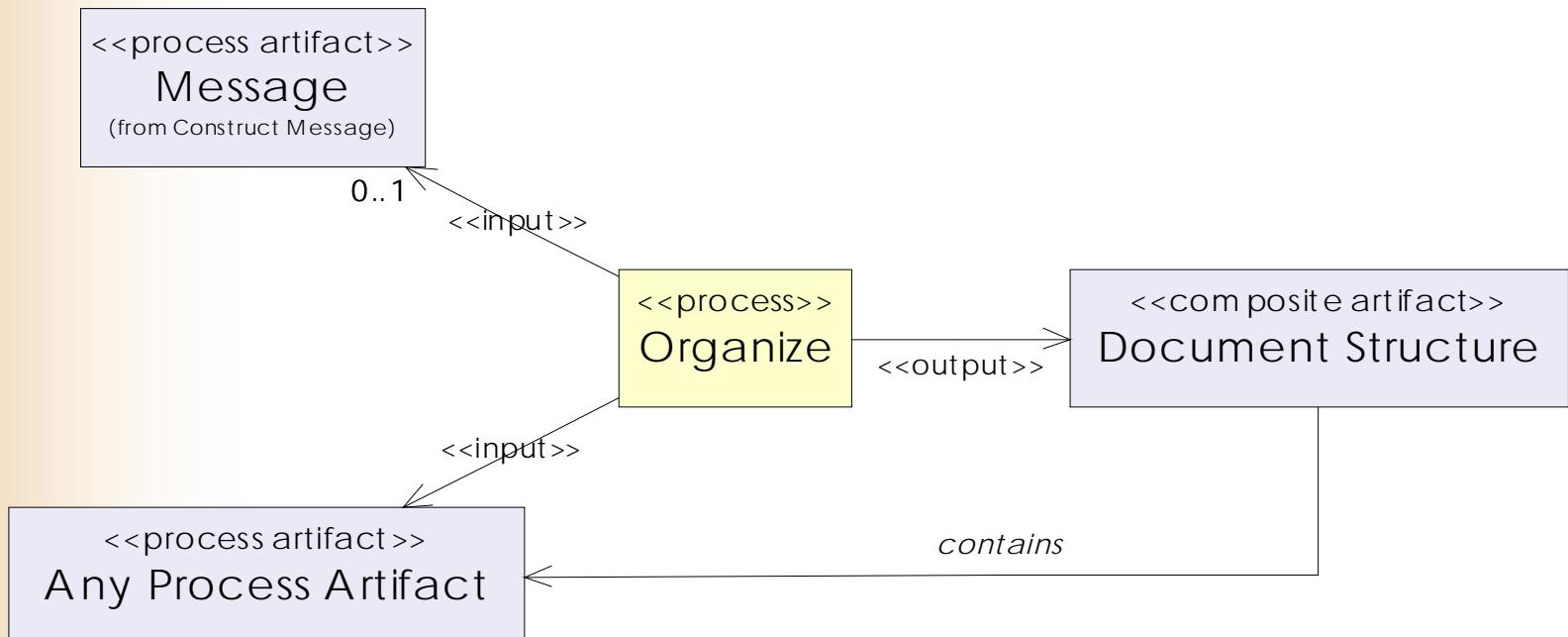
Publish

- Process where final content and user interface is created
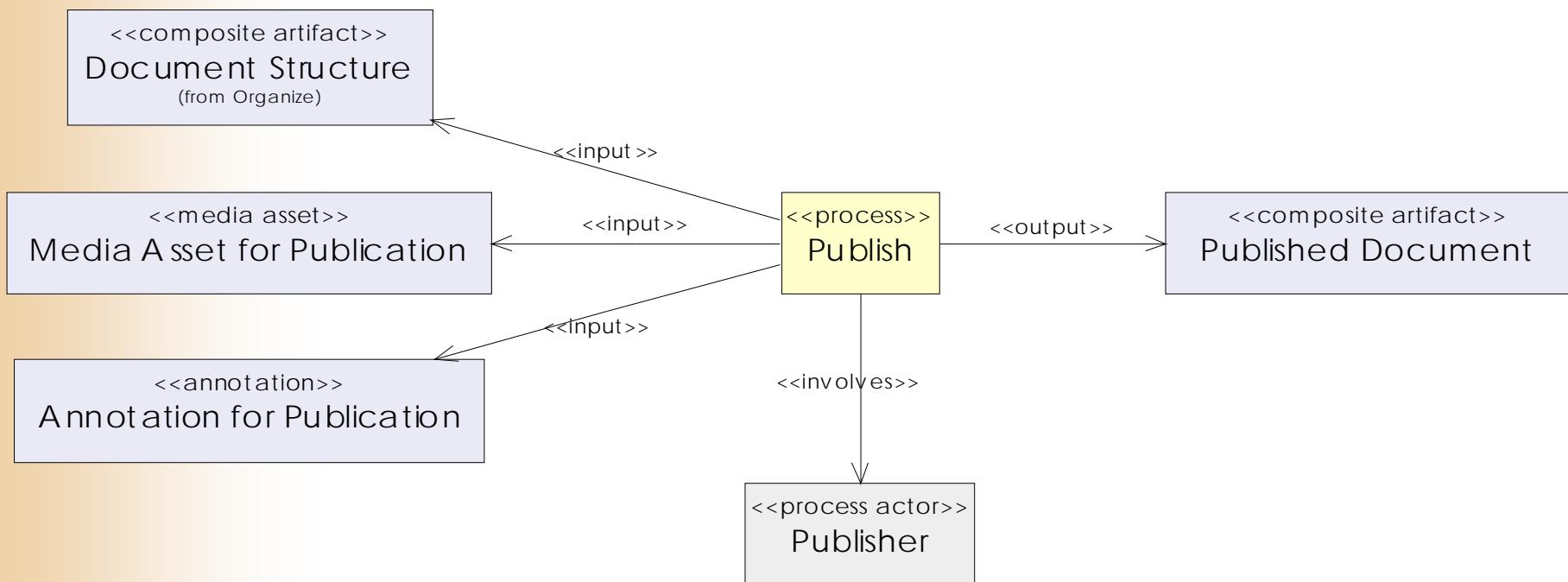
Distribute

- Process where final interaction between end-users and produced media occurs


```
┌──────────────────────┐                    ┌──────────────┐                    ┌──────────────────────────┐
│ <<composite artifact>>│  <<input>>         │  <<process>> │   <<output>>       │ <<external world artifact>>│
│  Published Document   │◄───────────────────│   Distribute │───────────────────►│    Real World Output      │
│    (from Publish)     │                    │              │                    │                           │
└──────────────────────┘                    └──────┬───────┘                    └──────────────────────────┘
                                                    │ <<involves>>
                                                    ▼
                                    ┌───────────────────────────────────┐
                                    │        <<process actor>>          │
                                    │ Distribution Hardware and Software │
                                    └───────────────────────────────────┘
```
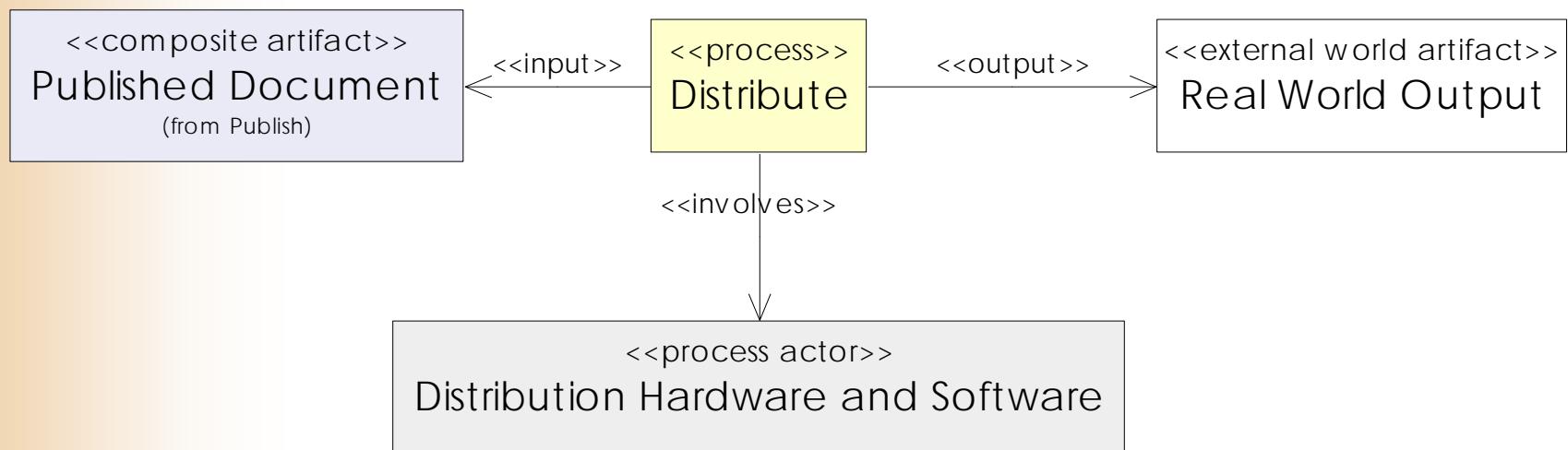
Canonical Processes Possible Flow
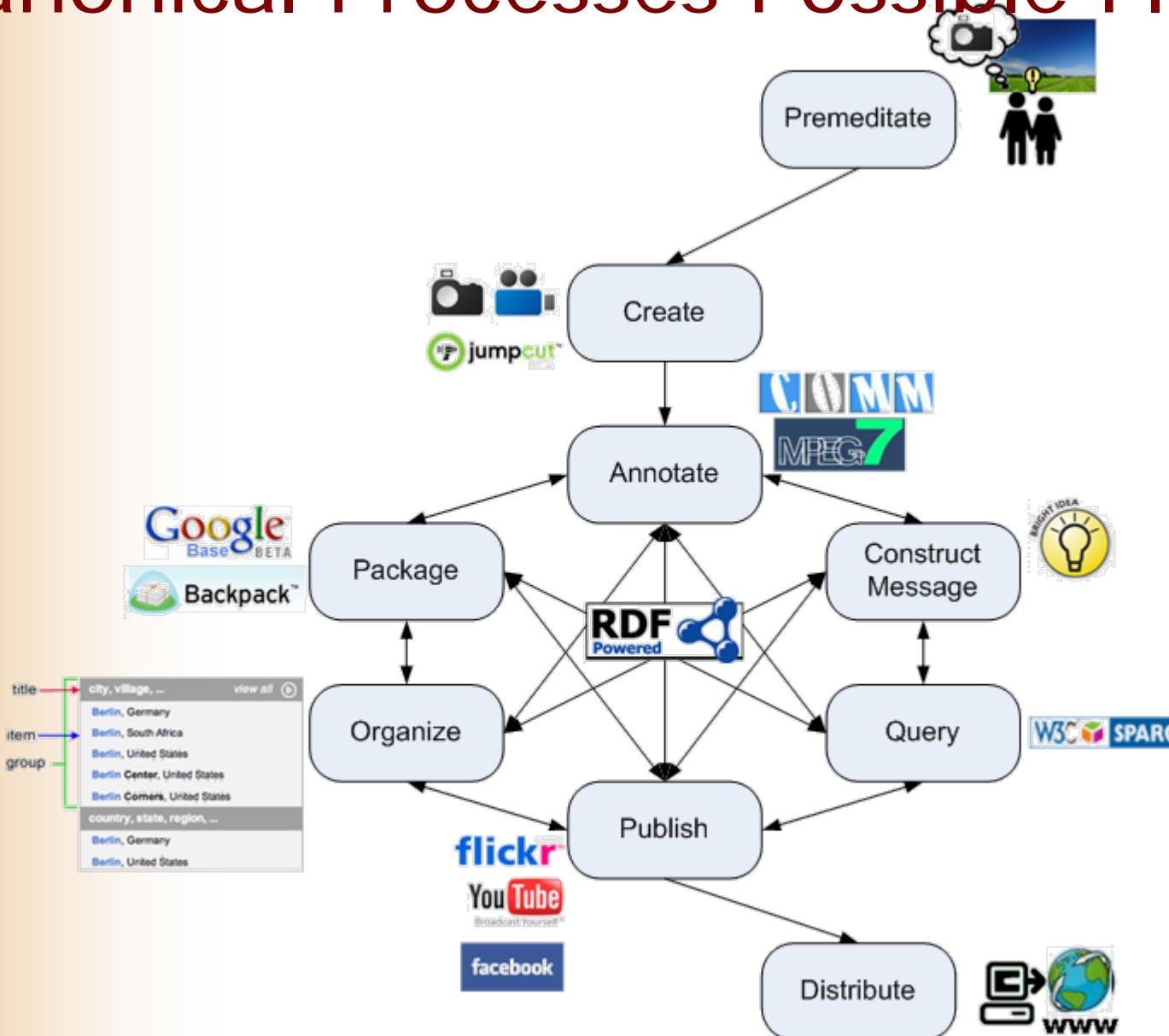
Sum Up

- Community agreement, not "yet another model"
- Large proportion of the functionality provided by multimedia applications can be described in terms of this model
- Initial step towards the definition of open web-based data structures for describing and sharing semantically annotated media assets

Discussion

- Frequently asked questions
    - Complex processes
    - Interaction
    - Complex artifacts and annotations can be annotated

- Towards a more rigorous formalization of model
    - Relationship to foundational ontologies
    - Semantics of Annotations

Upper Ontology

Model of Canonical Processes of Media Production

Models of Specific Media Production Processes

Concrete Systems

Literature

- Lynda Hardman: *Canonical Processes of Media Production*. In [Proceedings of the ACM Workshop on Multimedia for Human Communication - From Capture to Convey (MHC 05)](), November 2005.
- Special Issue on Canonical Processes of Media Production [http://www.ifi.uio.no/MMSJ/upcomming.html]() [http://www.cwi.nl/~media/projects/canonical/]()
- Lynda Hardman, Zeljko Obrenovic, Frank Nack, Brigitte Kerhervé and Kurt Piersol: *Canonical Processes of Semantically Annotated Media Production*. In [Multimedia Systems Journal](), 2008 (*to appear*)
- Philipp Sandhaus, Sabine Thieme and Susanne Boll: *Canonical Processes in Photo Book Production*. In [Multimedia Systems Journal](), 2008 (*to appear*)
- Stefano Bocconi, Frank Nack and Lynda Hardman: *Automatic generation of video documentaries*. In [Journal of Web Semantics](), 2008 (*to appear*).

Agenda

1. **Understanding Multimedia Applications Workflow**

   – CeWe Color Photo Book creation application

   – Vox Populi argumentative video sequences generation system

   – *The Canonical Processes of Media Production*

2. **Semantic Annotation of Multimedia Content**

   – Multimedia metadata formats: use cases and requirements

   – Multimedia metadata interoperability issues

   – MPEG-7 based ontologies

   – *COMM: A Core Ontology for MultiMedia*

3. **Semantic Search and Presentation of Multimedia Content**

   – Link your data!

   – *Searching and Browsing Multimedia Semantic Datasets with Cliopatria*

The Importance of the Annotations

EXIF.org

flickr GAMMA

Tagging

MPEG 7

You Tube
Broadcast Yourself™

Multimedia

last·fm

D 3v2

ifra

IPTC

xmp™
Adding Intelligence to Media

RDF Powered

FOAF + SKOS

Web

WWW

Semantic
Web

OWL
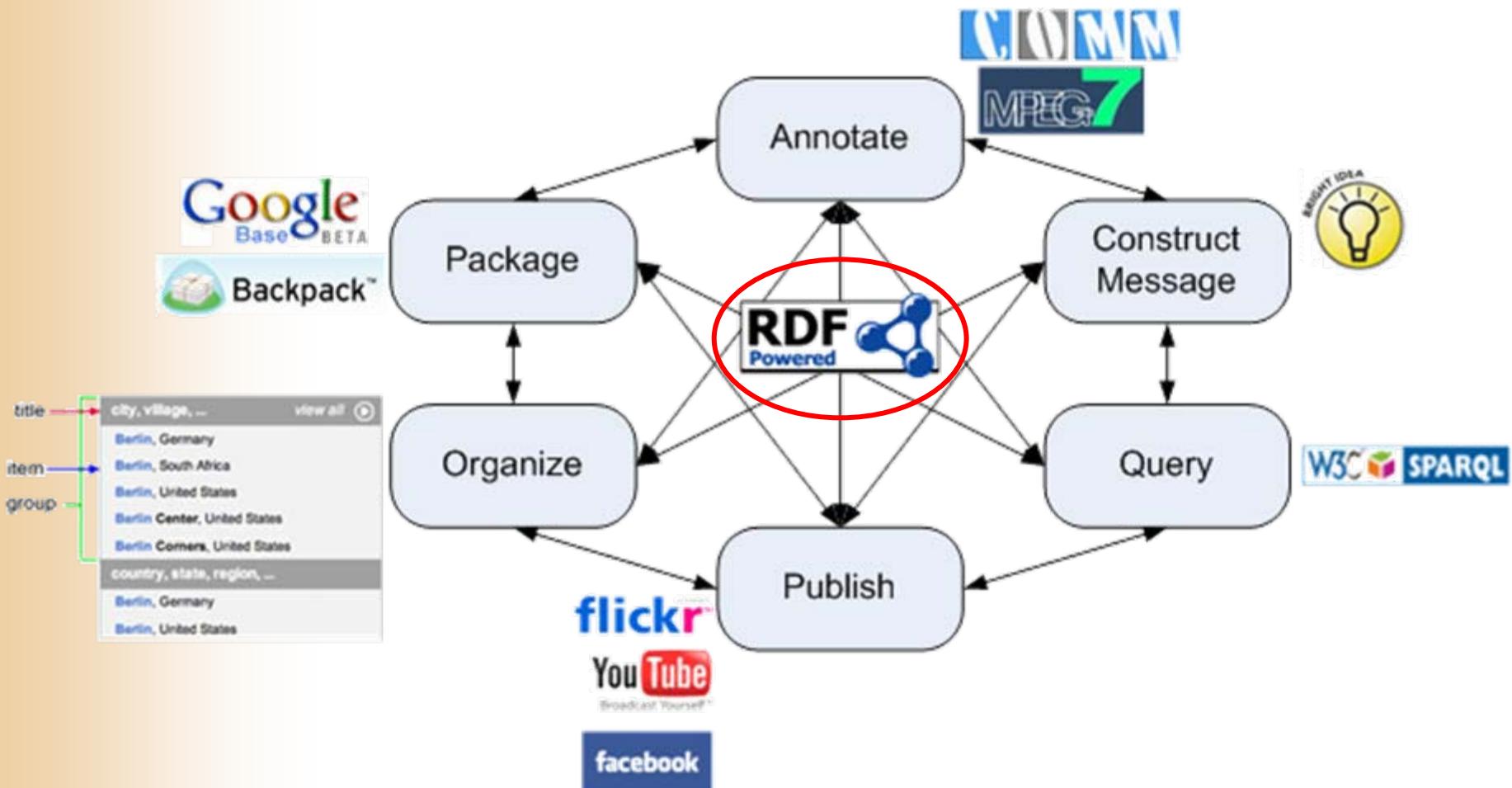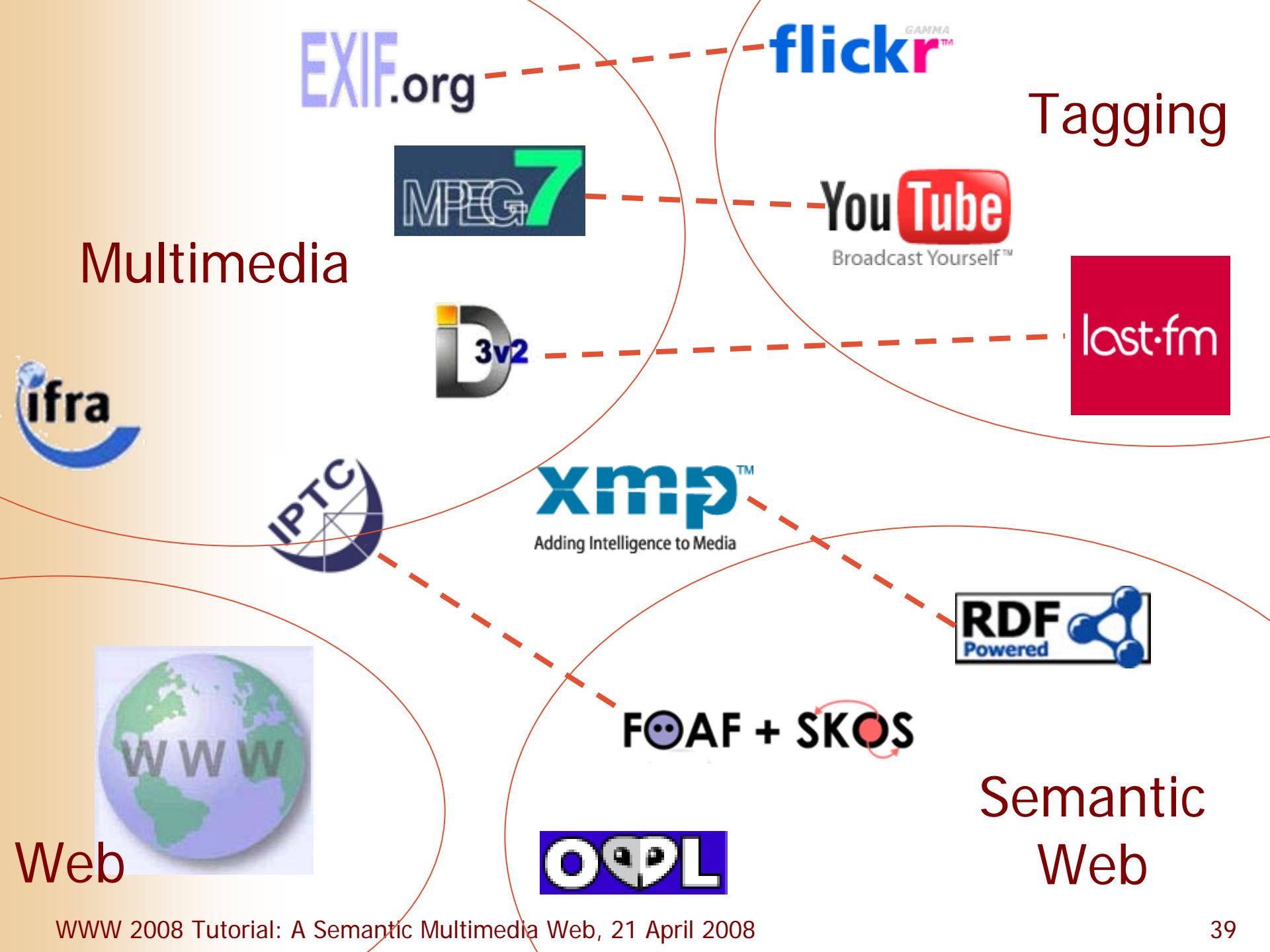
W3C Multimedia Semantics XG

W3C Multimedia Semantics XG
http://www.w3.org/2005/Incubator/mmsem/

Managing Personal Photos

- Interoperable Image Metadata
  - Combining EXIF, MPEG-7, IPTC and DIG35 metadata using RDF and OWL schemas

Facetting Music Songs

- Interoperable Music and Social Metadata
  - ID3 Tags + low-level features extraction + lastFM recommendations + FOAF profiles + …
  - Auto-construction of playlist (similar bit rate), Personalization, Browsing music store

Multimedia: Description methods

ISO

MPEG-21
MPEG-7
MPEG-4
MPEG-2
MPEG-1

W3C

User Interface & applications

Trust

Proof

Unifying Logic

Query: SPARQL | ontology: OWL | Rules: RIF

RDF-S

Data interchange: RDF

XML

URI | Unicode

Crypto
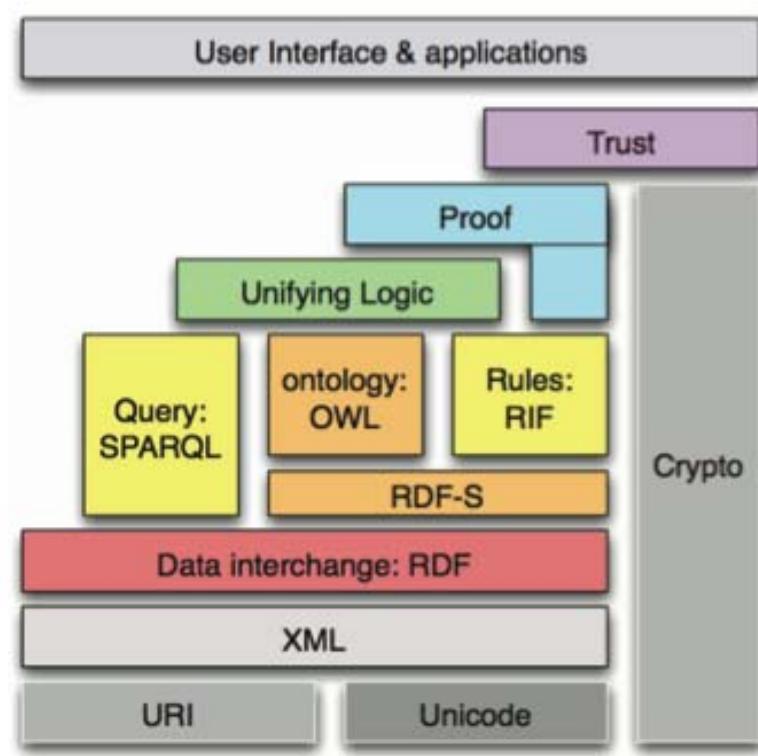
MPEG-7: a multimedia description language?

- ISO standard since December of 2001

- Main components:
  - Descriptors (Ds) and Description Schemes (DSs)
  - DDL (XML Schema + extensions)

- Concern all types of media

Content organization	Collections	Models	User interaction

- Creation & Production
- Media
- Usage
- Content management
- Content description
- Structural aspects
- Semantic aspects
- Navigation & Access
- Summaries
- Views
- Variations
- User Preferences
- User History

Basic elements
- Schema Tools
- Basic datatypes
- Links & media localization
- Basic Tools
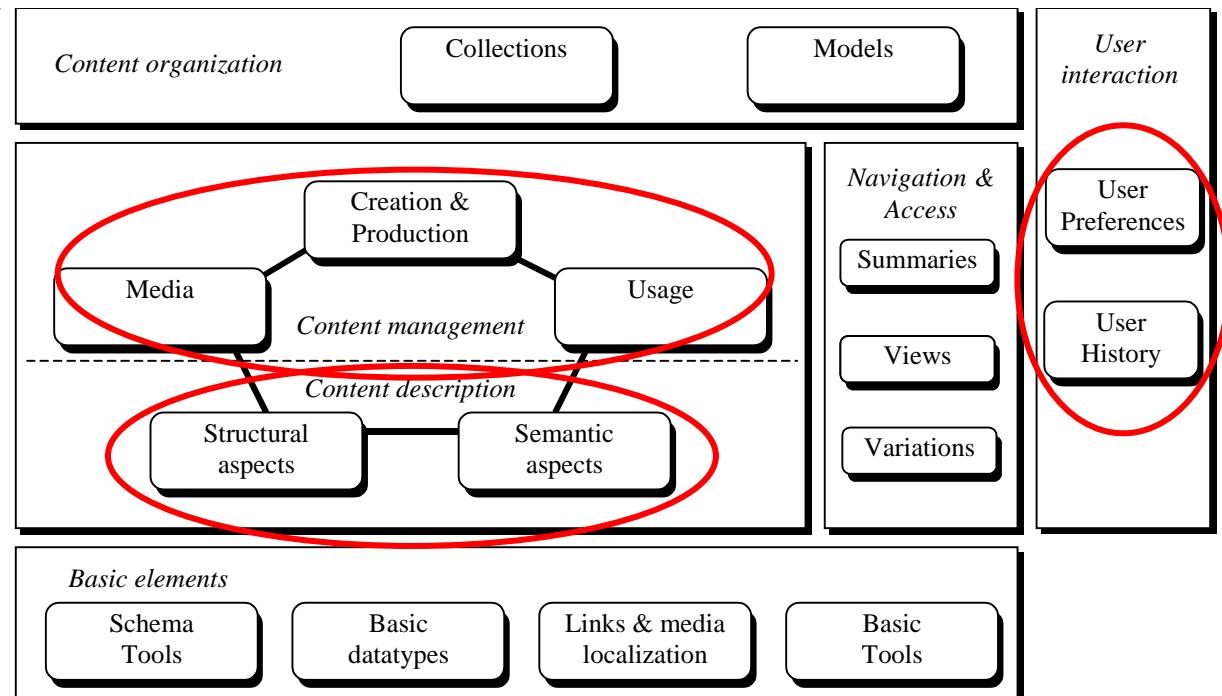
Part 5 – MDS
Multimedia Description Schemes

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

	Hunter	DS-MIRF	Rhizomik	COMM
Foundational Ontologies	ABC	None	None	DOLCE
Complexity	OWL-Full	OWL-DL	OWL-DL	OWL-DL
Coverage	MDS+Visual	MDS+CS	All	MDS+Visual
Applications	Digital Libraries	Digital Libraries	Digital Rights	MM Analysis

Common Scenario

The "Big Three" at the Yalta Conference (Wikipedia)

Common Scenario: Tagging Approach

Reg1

The "[Big Three](#)" at the Yalta Conference (Wikipedia)

- Localize a region
  - Draw a bounding box, a circle around a shape
- Annotate the content
  - Interpret the content
  - Tag: Winston Churchill, UK Prime Minister, Allied Forces, WWII

Common Scenario: SW Approach

**Reg1**

The "[Big Three](#)" at the Yalta Conference (Wikipedia)

- Localize a region
  - Draw a bounding box, a circle around a shape
- Annotate the content
  - Interpret the content
  - Link to knowledge on the Web

```
:Reg1 foaf:depicts dbpedia:WinstonChurchill
dbpedia:Churchill rdfs:label "Winston Churchill"
dbpedia:Churchill rdf:type foaf:Person
```

Hunter's MPEG-7 Ontology

http://en.wikipedia.org/wiki/Image:Yalta_Conference.jpg

mpeg7:MediaLocator

mpeg7:StillRegion

rdf:type

mpeg7:image

mpeg7:spatial_decomposition

Reg1

mpeg7:DominantColor

rgb(25,255,255)

mpeg7:depicts

The Big Three at the Yalta Conference

mpeg7:SpatialMask

mpeg7:depicts

mpeg7:Polygon

dbpedia:Churchill

mpeg7:Coords

5 25 10 20 15 15 10 10 5 15"^^xsd:string

DS-MIRF MPEG-7 Ontology

http://en.wikipedia.org/wiki/
Image:Yalta_Conference.jpg

mpeg7:MediaURI

mpeg7:MediaLocator

mpeg7:StillRegion

rdf:type

mpeg7:image

mpeg7:SpatialDecomposition

Reg1

dbpedia:Churchill

mpeg7:RelatedMaterial

mpeg7:CreationInformation

mpeg7:SpatialMask

mpeg7:Creation

mpeg7:SubRegion

mpeg7:Polygon

mpeg7:Coords

mpeg7:Title

mpeg7:dim

The Big Three at the Yalta
Conference

contentString

5 25 10 20 15 15 10 10 5 15"^^xsd:string

Rhizomik MPEG-7 Ontology

http://en.wikipedia.org/wiki/
Image:Yalta_Conference.jpg

mpeg7:MediaLocator

mpeg7:SegmentType

mpeg7:image

mpeg7:spatial_decomposition

rdf:type

Reg1

dbpedia:Churchill

mpeg7:Semantic

mpeg7:CreationInformation

mpeg7:SpatialMask

mpeg7:SubRegion

mpeg7:Polygon

mpeg7:Coords

mpeg7:Title

mpeg7:dim

The Big Three at the Yalta
Conference

5 25 10 20 15 15 10 10 5 15"^^xsd:string

COMM: Fragment Identification

http://en.wikipedia.org/wiki/
Image:Yalta_Conference.jpg

dns:realized-by

core:image-data

dns:setting

core:semantic-annotation

dns:plays

dns:defines

foaf:Person

loc:region-locator-descriptor

loc:spatial-mask-role

dns:played-by

core:semantic-label-role

dns:defines

dns:played-by

rdf:type

loc:bounding-box

5 25 10 20 15 15 10 10 5 15"^^xsd:string

dbpedia:Churchill

data:has-rectangle

Comparison

- Link with domain semantics
    - Hunter: ABC model + `mpeg7:depicts` relationship
    - DS-MIRF: Domain ontologies needs to subclass the general MPEG-7 categories
    - Rhizomik: Use the `mpeg7:semantic` relationship
    - COMM: Semantic Annotation pattern
- MPEG-7 coverage
    - Hunter: extension of the MPEG-7 visual descriptors
    - COMM:
        - Formalization of the context of the annotation
        - Representation of the method (algorithm) that provides the annotation

Comparison

- 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:

	Hunter	DS-MIRF	Rhizomik	COMM
Triples	11	27	20	19

Scenario: Image

Reg1

The "[Big Three]" at the Yalta Conference (Wikipedia)

- Localize a region (bounding box)
- Annotate the content (interpretation)
  - Tag: Winston Churchill, UK Prime Minister, Allied Forces, WWII
  - Link to knowledge on the Web

```
:Reg1 foaf:depicts dbpedia:WinstonChurchill
dbpedia:Churchill rdfs:label "Winston Churchill"
dbpedia:Churchill rdf:type foaf:Person
```

Scenario: Video

Seq4

Seq1

A history of G8 violence ([video](#))
(© Reuters)

- Localize a region
- Annotate the content
  - Tag: G8 Summit, Heiligendamn, 2007
  - Link to knowledge on the Web

EU Summit, Gothenburg, 2001

```
:Seq1 foaf:depicts dbpedia:34th_G8_Summit
:Seq4 foaf:depicts dbpedia:EU_Summit
geo:Heilegendamn skos:broader geo:Germany
```

Research Problem

Reg1

Seq1

Seq4

The "Big Three" at the Yalta Conference (Wikipedia)

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

- Multimedia objects are complex
  - Compound information objects, fragment identification   ⇨ **MPEG-7**
- Semantic annotation
  - Subjective interpretation, context dependent   ⇨ **D&S | OIO**
- Linked data principle
  - Open to reuse existing knowledge   ⇨ **RDF**

COMM: Design Rationale

- Approach:
  - NO 1-to-1 translation from MPEG-7 to OWL/RDF
  - Need for patterns: use DOLCE, a well designed foundational ontology as a modeling basis

- Design patterns:
  - Ontology of Information Objects (OIO)
    - Formalization of information exchange
    - Multimedia = complex compound information objects
  - Descriptions and Situations (D&S)
    - Formalization of context
    - Multimedia = contextual interpretation (situation)

- Define **multimedia patterns** that translate MPEG-7 in the DOLCE vocabulary

COMM: Core Functionalities

- Most important MPEG-7 functionalities:
  - **Decomposition** of multimedia content into segments
  - **Annotation** of segments with metadata
    - Administrative metadata: creation & production
    - Content-based metadata: audio/visual descriptors
    - Semantic metadata: interface with domain specific ontologies

⇨ Note that all are subjective and context dependent situations

COMM: D&S / OIO Patterns

D&S / OIO

description | information-object | role | method | situation
structured-data-description | digital-data | processing-role | algorithm
descriptor | multimedia-data | output-role | input-role
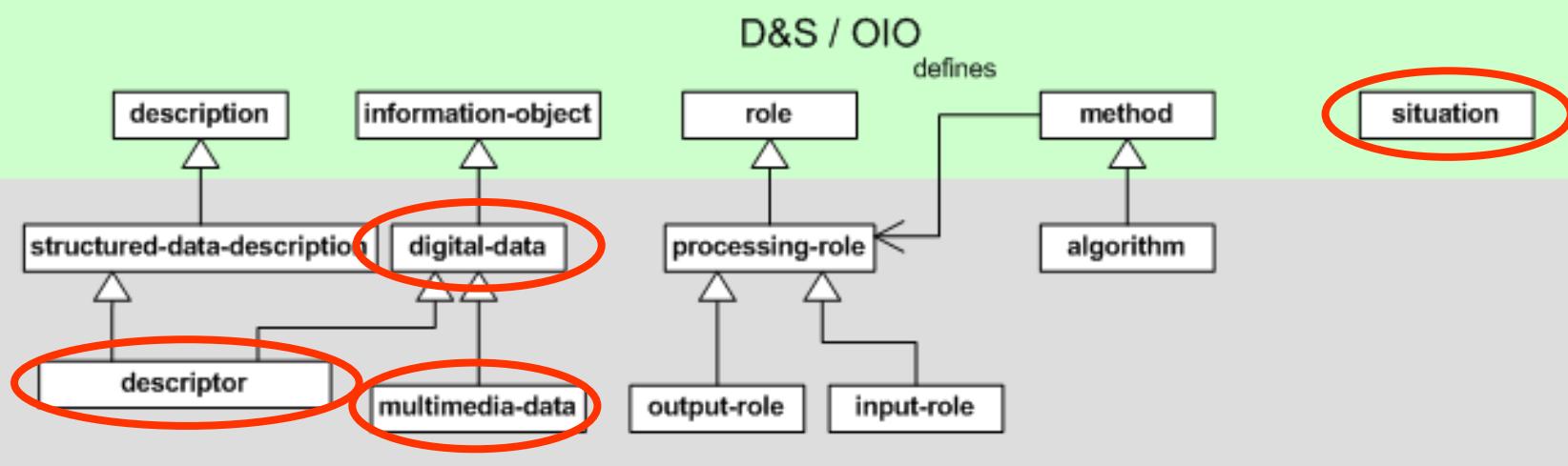
■Definition of design patterns for **decomposition** and **annotation** based on D&S and OIO

- MPEG-7 describes digital data (*multimedia information objects*) with digital data (*annotation*)
- *Digital data* entities are information objects
- Decompositions and annotations are *situations* that satisfy the rules of a method or algorithm

COMM: Decomposition Pattern
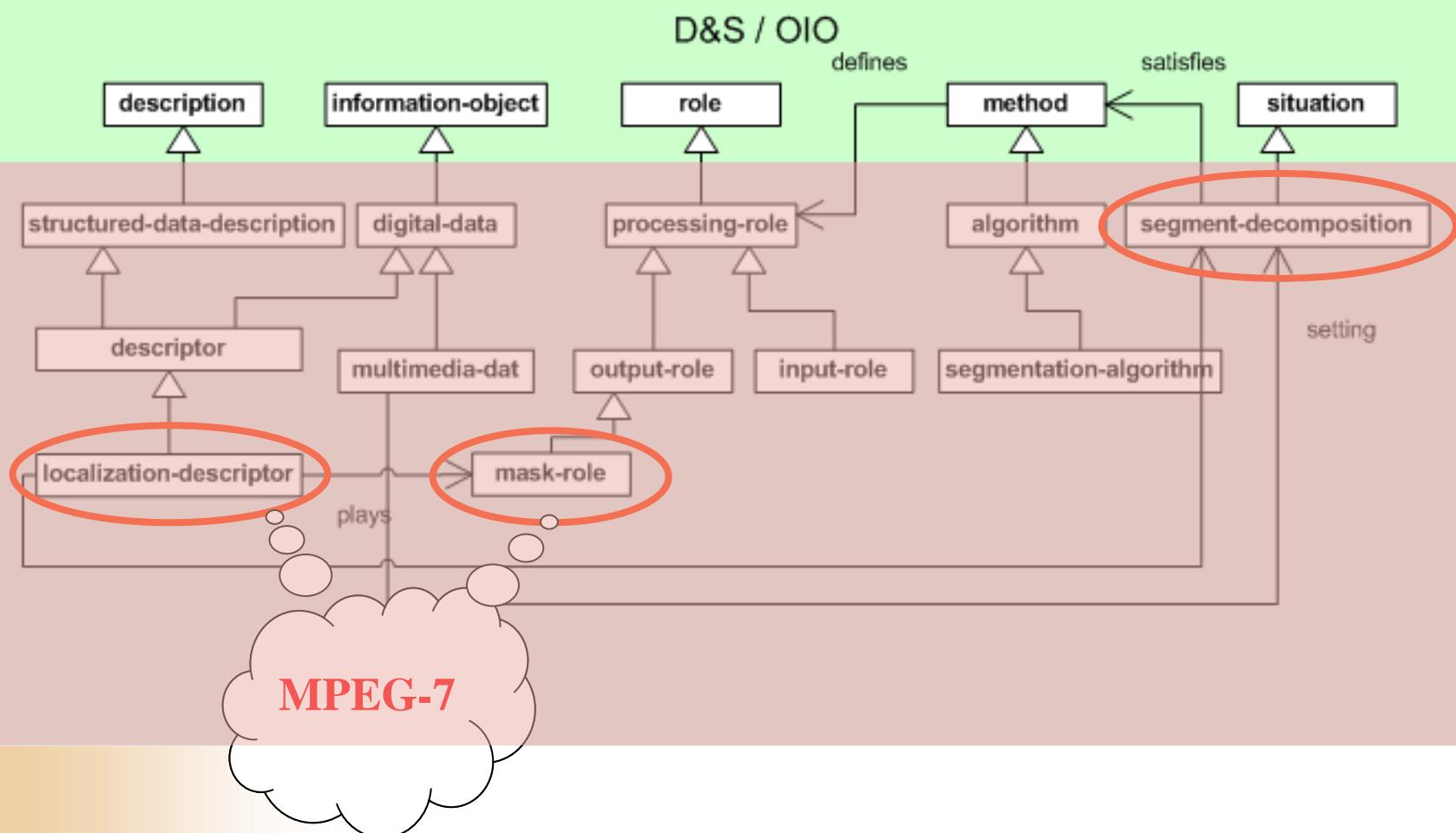
COMM: Annotation Pattern
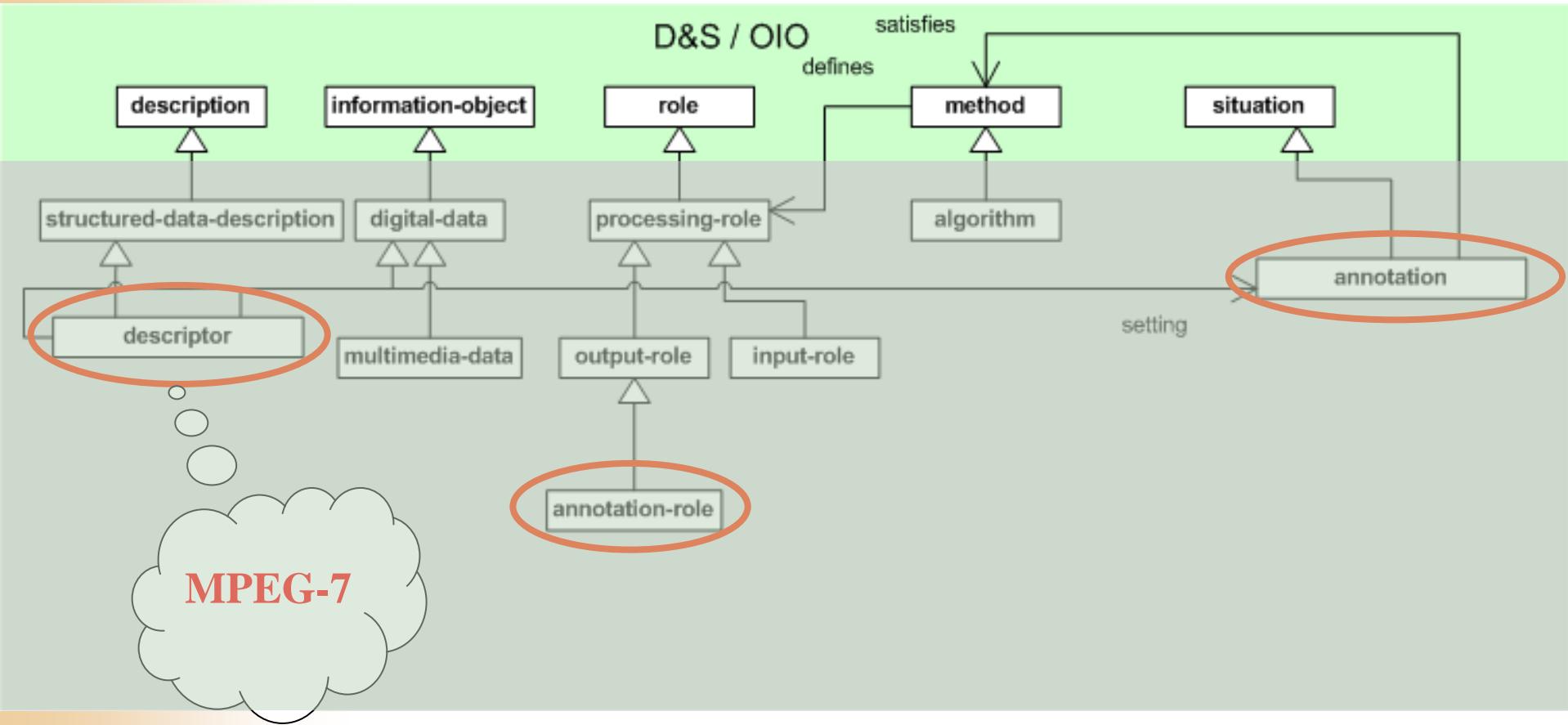
COMM: Semantic Pattern
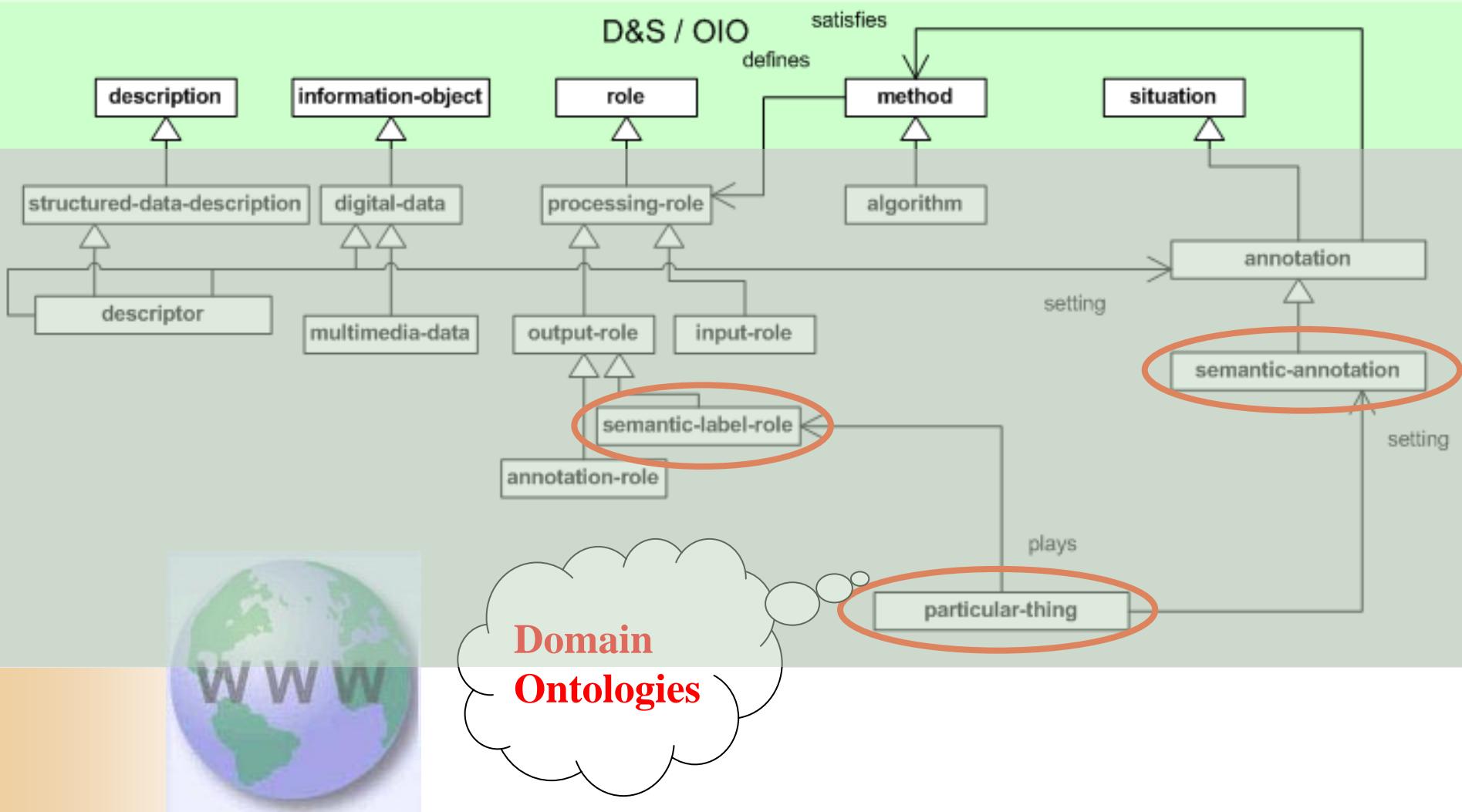
# COMM: Modules

Annotation Pattern

Decomposition Pattern

Fundamental Knowledge about the World

Knowledge about a specific Domain

DOLCE

Connected by Semantic Annotation Pattern

Descriptions & Situation

Ontology of Information Objects

Domain Ontolog

Datatype

Core

Localization

Media

Visual

Audio

Text / LingInfo

Multimedia Knowledge (COMM)
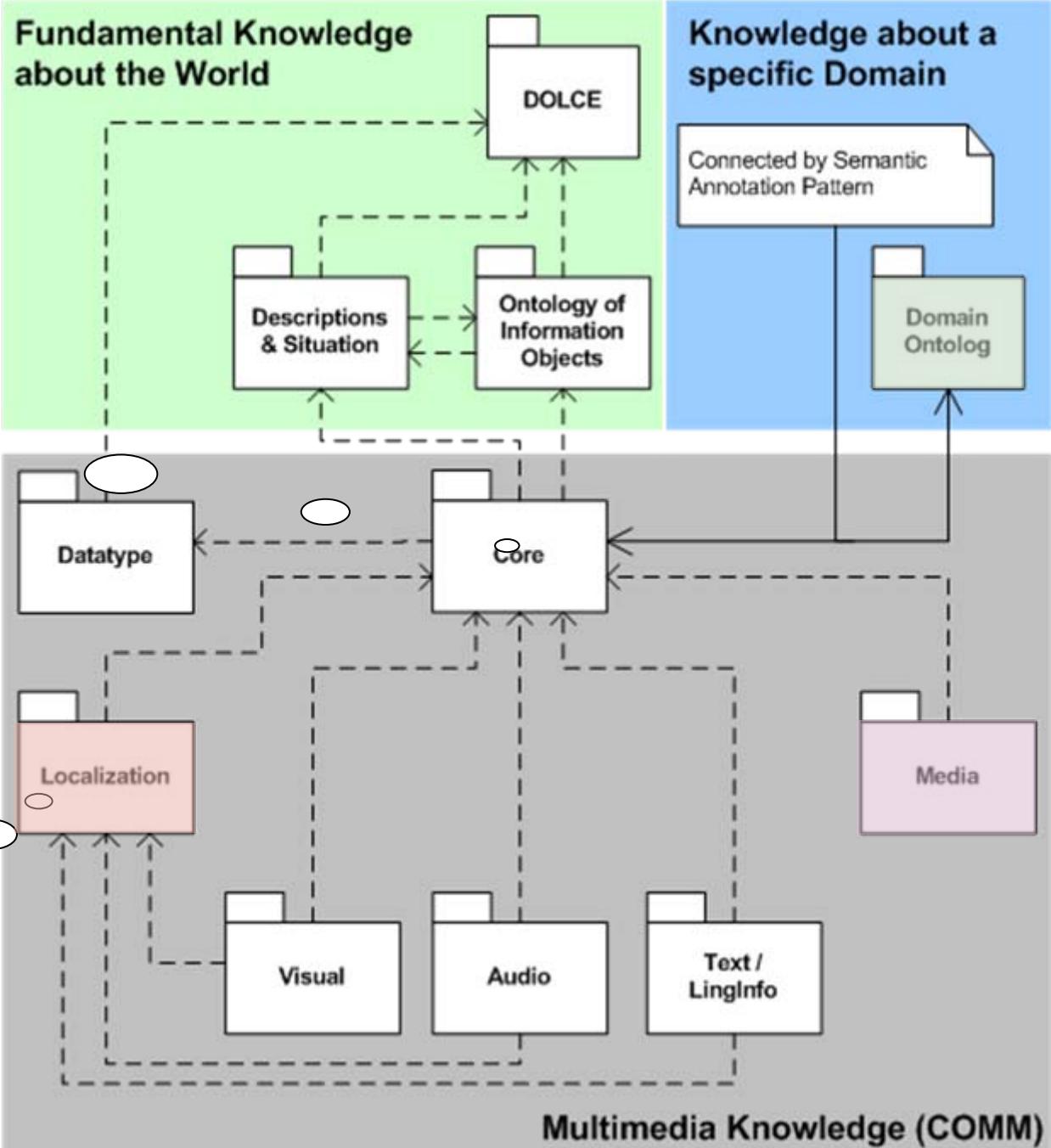
Example 1: Fragment Identification

http://en.wikipedia.org/wiki/
Image:Yalta_Conference.jpg

dns:realized-by

core:image-data

dns:plays

loc:region-
locator-descriptor

loc:spatial-mask-
role

dns:played-by

dns:defines

loc:bounding-box

5 25 10 20 15 15 10 10 5 15"^^xsd:string

data:has-rectangle
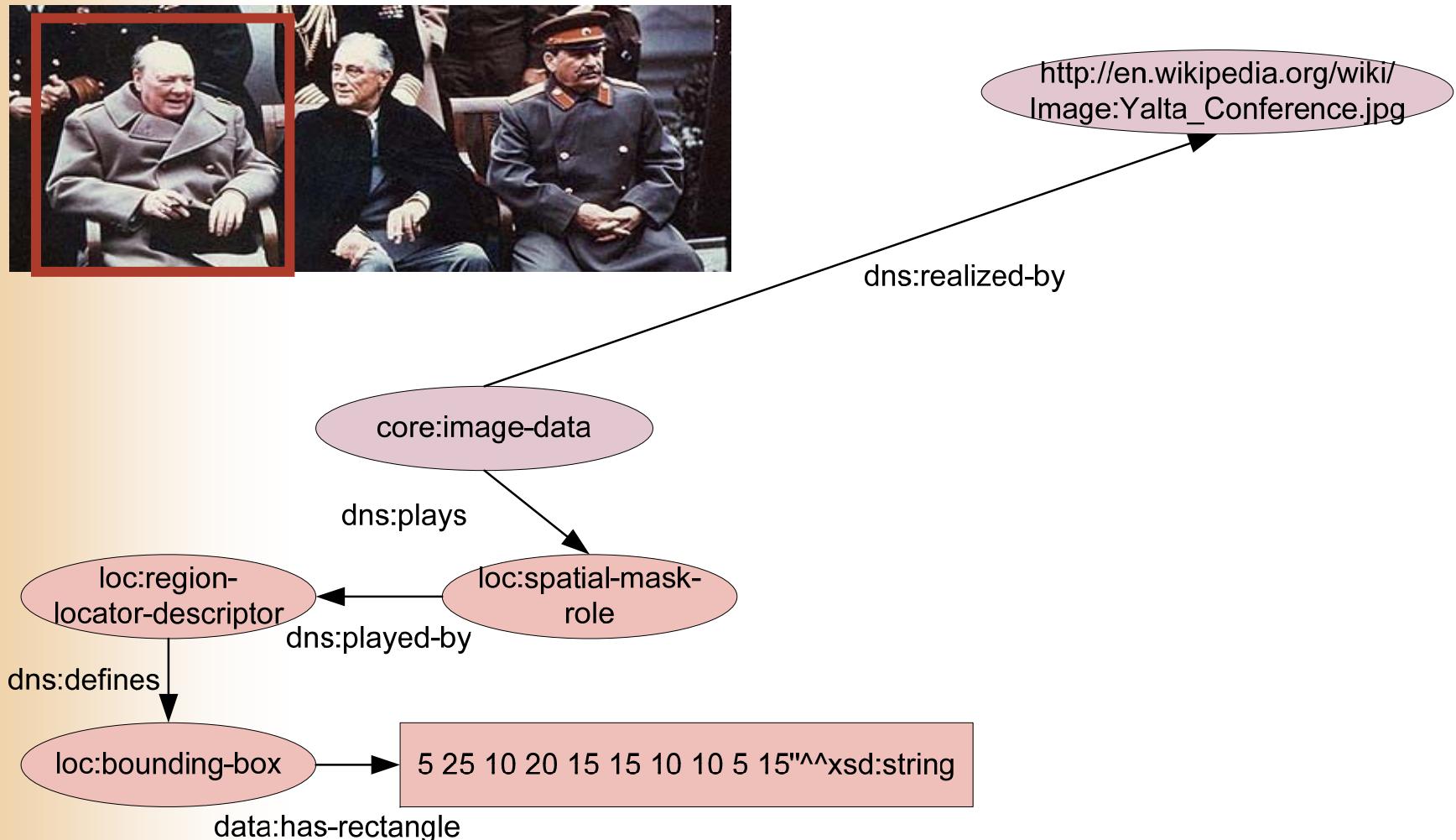
Example 1: Region Annotation

- http://en.wikipedia.org/wiki/Image:Yalta_Conference.jpg

dns:realized-by

core:image-data → dns:setting → core:semantic-annotation

dns:plays

core:image-data → loc:spatial-mask-role

dns:defines

core:semantic-annotation → core:semantic-label-role

foaf:Person

loc:region-locator-descriptor ← dns:played-by ← loc:spatial-mask-role

dns:defines

loc:bounding-box → data:has-rectangle → 5 25 10 20 15 15 10 10 5 15"^^xsd:string

core:semantic-label-role → dns:played-by → http://en.wikipedia.org/wiki/Churchill

rdf:type → foaf:Person

Example 2: Fragment Identification

http://www.reuters.com/news/video/
summitVideo?videoId=56114

dns:realized-by

core:image-data

dns:plays

loc:media-time-descriptor

dns:played-by

loc:temporal-mask-role

dns:defines

loc:media-time-point

data:has-time

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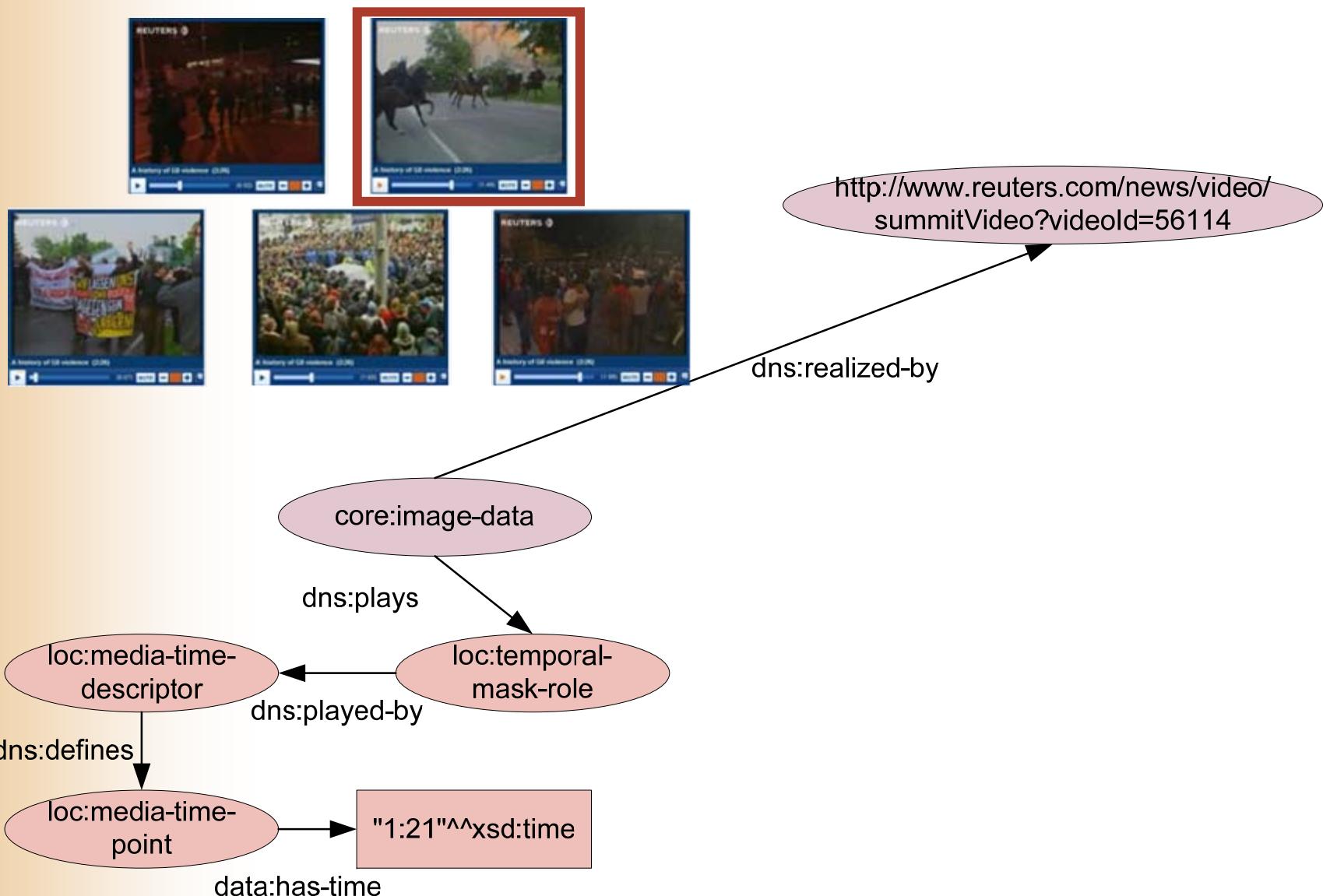
Example 2: Sequence Annotation

http://www.reuters.com/news/video/summitVideo?videoId=56114

core:image-data

dns:realized-by

dns:setting → core:semantic-annotation

dns:plays

loc:media-time-descriptor

loc:temporal-mask-role

dns:played-by

dns:defines

loc:media-time-point

"1:21"^^xsd:time

data:has-time

dns:defines

core:semantic-label-role

tgn:Sweden

dns:played-by

skos:broader

tgn:Gothenburg
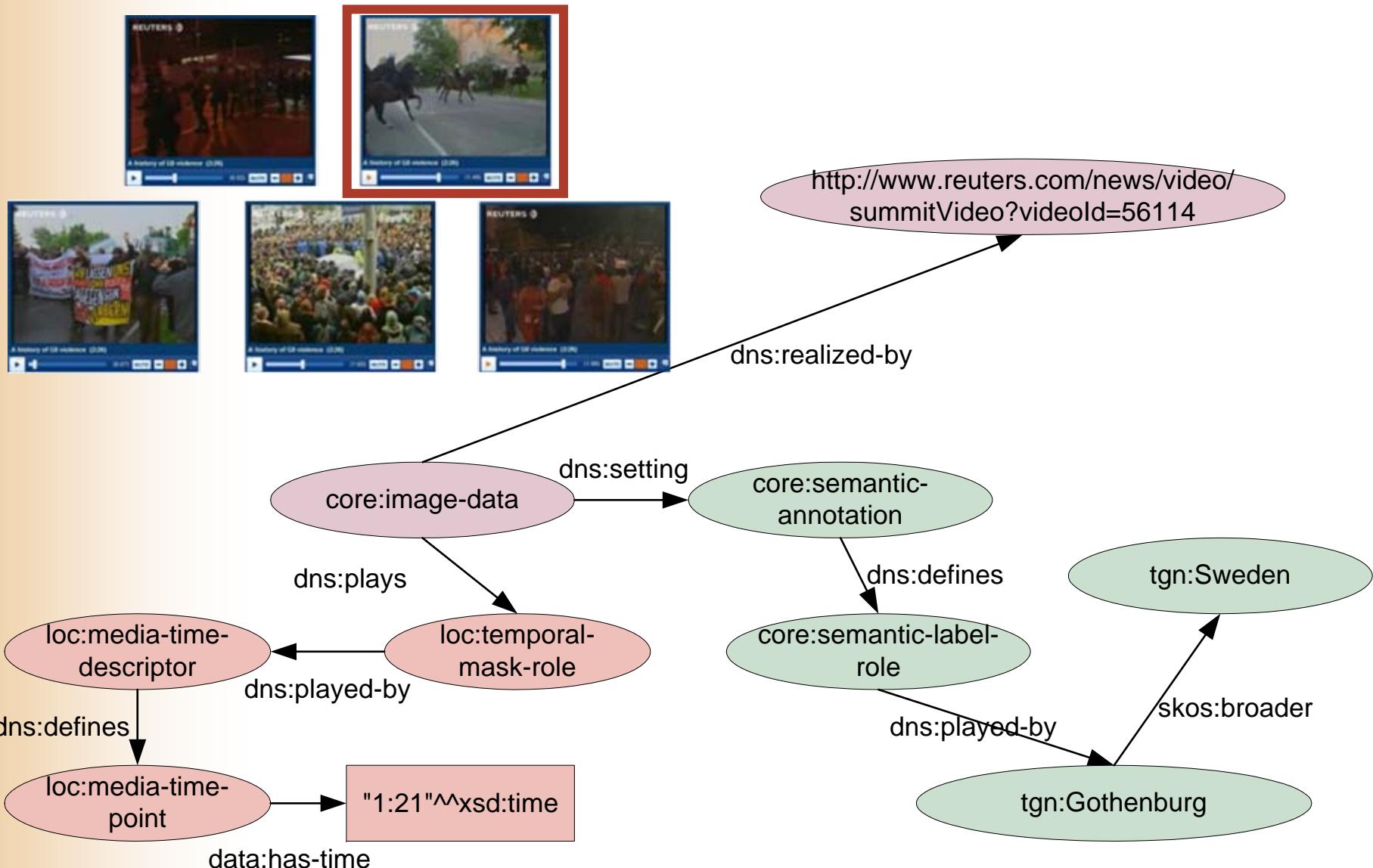
Implementation

- COMM fully formalized in OWL DL
  - Rich axiomatization, consistency check (Fact++v1.1.5)
  - OWL 2.0: 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

KAT Annotation Tool
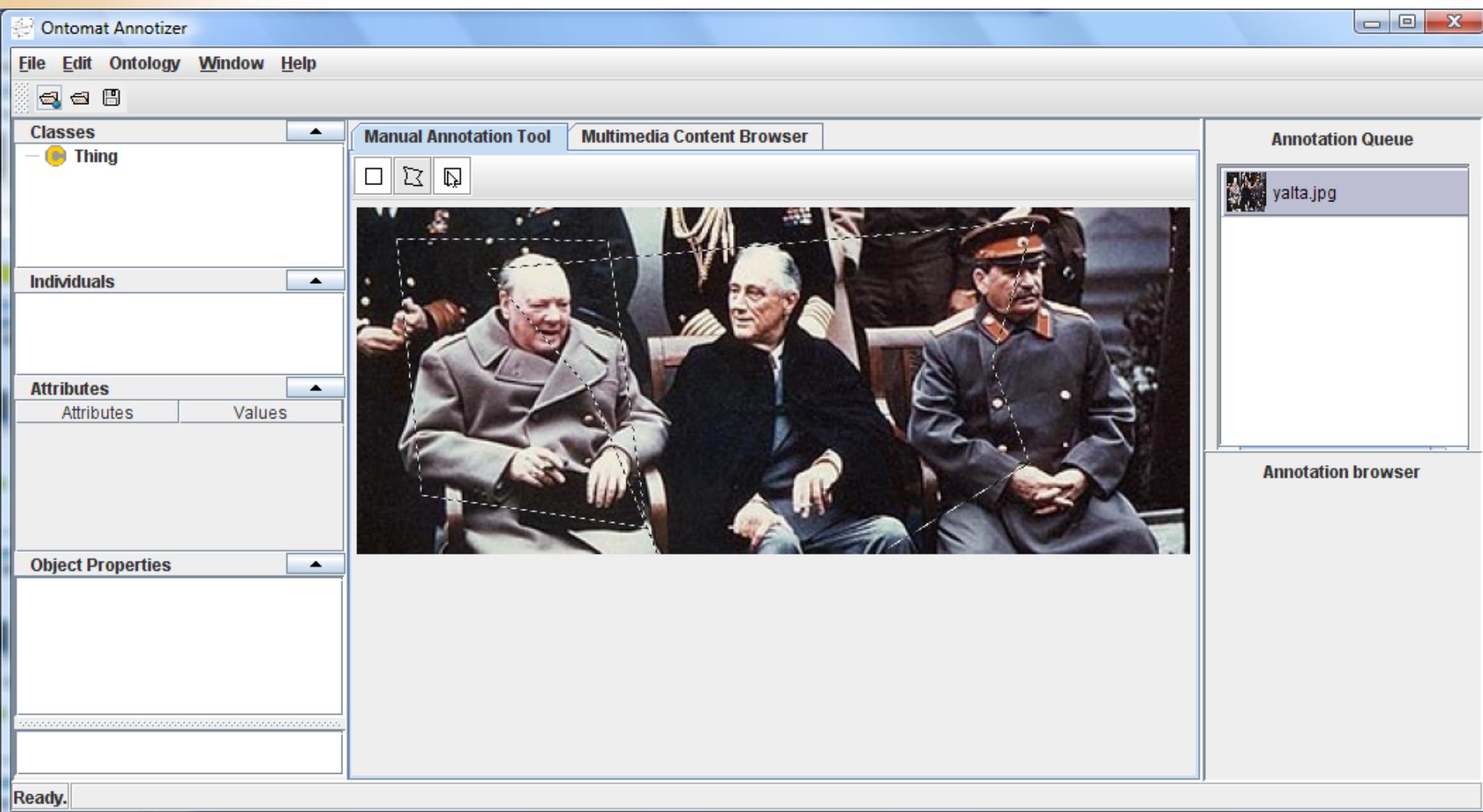
Evaluation

- Applied Domains
  - Knowledge management for multimedia documents
  - Driving multimedia analysis process
  - Generate new interfaces for browsing multimedia content
- Scalability
  - 4 minutes video,
    TRECVid metadata expressed in COMM
    - 250 K statements
  - Reasoning in large scale applications

Literature

- Michael Hausenblas *et al.*: [Multimedia Vocabularies on the Semantic Web](#). W3C Multimedia Semantics Incubator Group Report (XGR), 24 July 2007.

- Raphaël Troncy, Jacco van Ossenbruggen, Jeff Z. Pan and Giorgos Stamou. [Image Annotation on the Semantic Web](#). W3C Multimedia Semantics Incubator Group Report (XGR), 14 August 2007.

- Vassilis Tzouvaras, Raphaël Troncy and Jeff Z. Pan. [Multimedia Annotation Interoperability Framework](#). W3C Multimedia Semantics Incubator Group Report Editor's Draft, 14 August 2007.

- Richard Arndt, Raphaël Troncy, Steffen Staab, Lynda Hardman and Miroslav Vacura: *COMM: Designing a Well-Founded Multimedia Ontology for the Web*. In [6th International Semantic Web Conference (ISWC'2007)](#), Busan, Korea, November 11-15, 2007.

- Raphaël Troncy, Oscar Celma, Suzanne Little, Roberto Garcia, Chrisa Tsinaraki: *MPEG-7 based Multimedia Ontologies: Interoperability Support or Interoperability Issue?* In [1st Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies (MAReSO'2007)](#), Genoa, Italy, December 2007.

Agenda

1. Understanding Multimedia Applications Workflow
   - CeWe Color Photo Book creation application
   - Vox Populi argumentative video sequences generation system
   - *The Canonical Processes of Media Production*

2. Semantic Annotation of Multimedia Content
   - Multimedia metadata formats: use cases and requirements
   - Multimedia metadata interoperability issues
   - MPEG-7 based ontologies
   - *COMM: A Core Ontology for MultiMedia*

3. Semantic Search and Presentation of Multimedia Content
   - Link your data!
   - *Searching and Browsing Multimedia Semantic Datasets with Cliopatria*

A Giant Graph Open to the World


```
<rdf:Description
 rdf:about="Ganesh.jpg">
    <dc:title>An image of the
Elephant Ganesh</dc:title>
    <dc:creator>
Raphaël Troncy</dc:creator>
</rdf:Description>
```

- Annotate the content (interpretation)
  Elephant, Ganesh, Thailande, Holidays, Chiang Mai

- Link to knowledge on the Web
  ```
  :img foaf:depicts dbpedia:Ganesh
  dbpedia:Ganesh rdfs:label "Vinayaka"
  dbpedia:Ganesh skos:altlabel "Ganapati"
  dbpedia:Ganesh rdf:type wn:synset-Deities-noun-1
  dbpedia:Ganesh owl:sameas wn:synset-Ganesh-noun-1
  ```

Linking Open Data Project

- Expose open datasets in RDF
- Set RDF links among the data items for different datasets
- Over 2 billion triples, 3 millions links (March 2008)

http://richard.cyganiak.de/2007/10/lod/
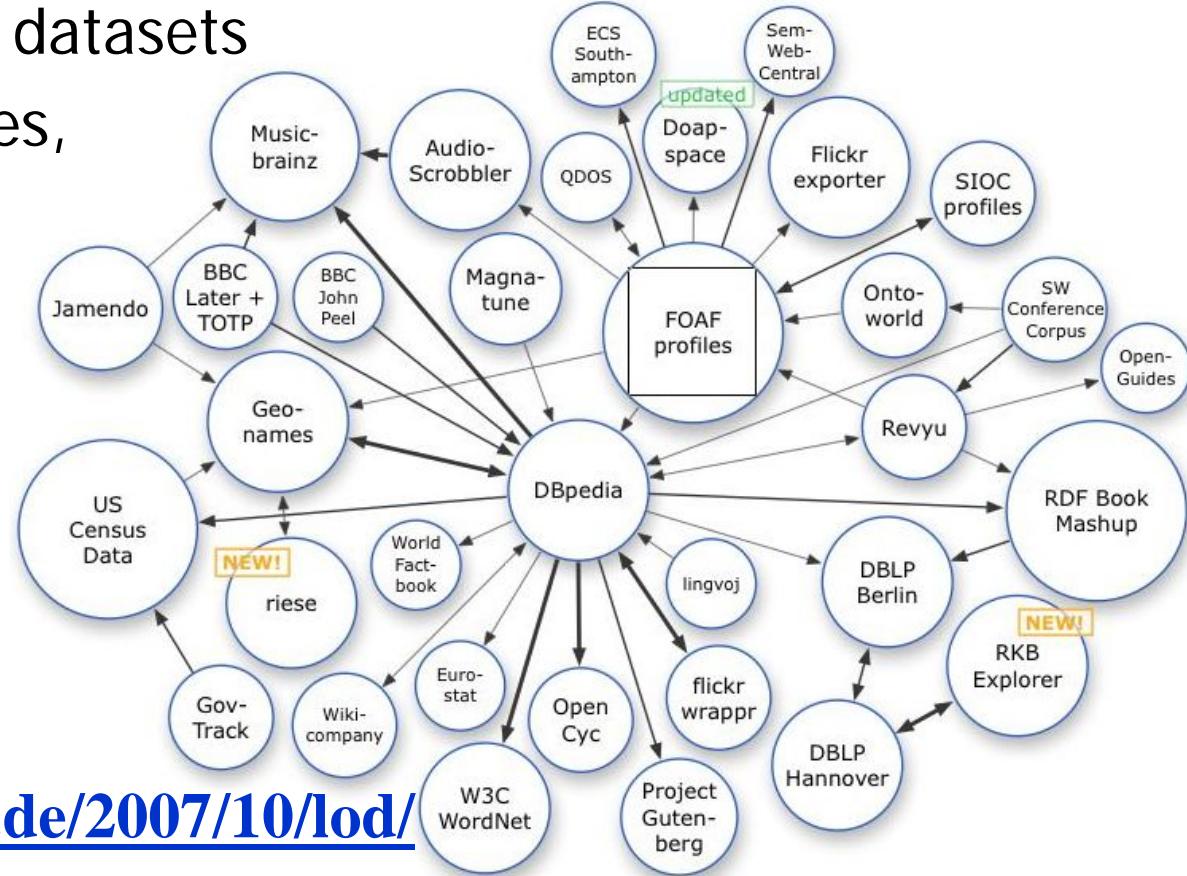
DBpedia

- DBpedia is a community effort to:
  - extract structured "infobox" information from Wikipedia
  - interlink DBpedia with other datasets on the Web
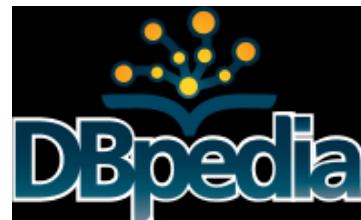
DBpedia

Extracting Infobox Data

```
http://en.wikipedia.org/wiki/Calgary
```

```
<http://dbpedia.org/resource/Calgary>
  dbpedia:native_name "Calgary" ;
  dbpedia:altitude "1048" ;
  dbpedia:population_city "988193" ;
  dbpedia:population_metro "1079310" ;
  mayor_name
        dbpedia:Dave_Bronconnier ;
  governing_body
        dbpedia:Calgary_City_Council ;
  . . .
```

■ **Altogether 9,100,000 RDF triples extracted from 754,000 infoboxes**

Calgary	
Downtown Calgary.	
Government	
- Mayor	Dave Bronconnier (Past mayors)
-	Calgary City Council
Governing body	
- Manager	Owen A. Tobert
Area [1]	
- City	726.50 km² (280.5 sq mi)
- Metro	5,107.43 km² (1,972 sq mi)
Elevation	1,048 m (3,438.3 ft)
Population (2006)[1]	
- City	988,193
- Density	1,360.2/km² (3,522.9/sq mi)
- Metro	1,079,310
- Population rank	3rd
- Metro rank	5th

Christian Bizer et al: DBpedia – Querying Wikipedia Like a Database (May 11, 2007)

Automatic Links Among Open Datasets

```
<http://dbpedia.org/resource/Calgary>
  owl.sameAs <http://sws.geonames.org/5913490>;

...
```

DBpedia
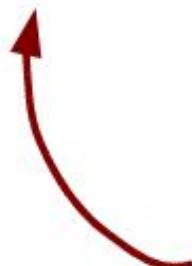
Geonames

```
<http://sws.geonames.org/5913490>
  owl:sameAs <http://DBpedia.org/resource/Calgary>
  wgs84_pos:lat "51.050112282";
  wgs84_pos:long "-114.085285152";
  sws:population "968460"

...
```

Processors can switch automatically from one to the other …

Take Home Message

- Reuse what is there

  – Of course, one could create RDF data manually ...
     ... but that is unrealistic on a large scale

  – Goal is to generate RDF data automatically when possible and "fill in" by hand only when necessary

    - service to get RDF from flickr images
      http://www.kanzaki.com/works/2005/imgdsc/flickr2rdf

    - service to get RDF from XMP
      http://www.ivan-herman.net/cgi-bin/blosxom.cgi/WorkRelated/SemanticWeb/xmpextract.html

- Expose what you make

Semantic Browsing of Multimedia News

- Goal:
  - Provide en environment for *searching* and *browsing* *contextualized multimedia news* information

- Method:
  - Semantic processing of multimedia news items
  - Link news items with knowledge on the web

- Datasets:
  - News stories: Jun/Jul 2006 (en/fr) newsfeed, AFP ± 90,000 items
  - Photos: 2006 football world cup, AFP ± 2,500 items
 - Video: Jun/Jul 2006 TV News (fr), INA ± 30 items

Problems

- No integration of media (stories, photo, video)
- Little (or no) context in the news presentation
- Lack of interoperability in the current workflow
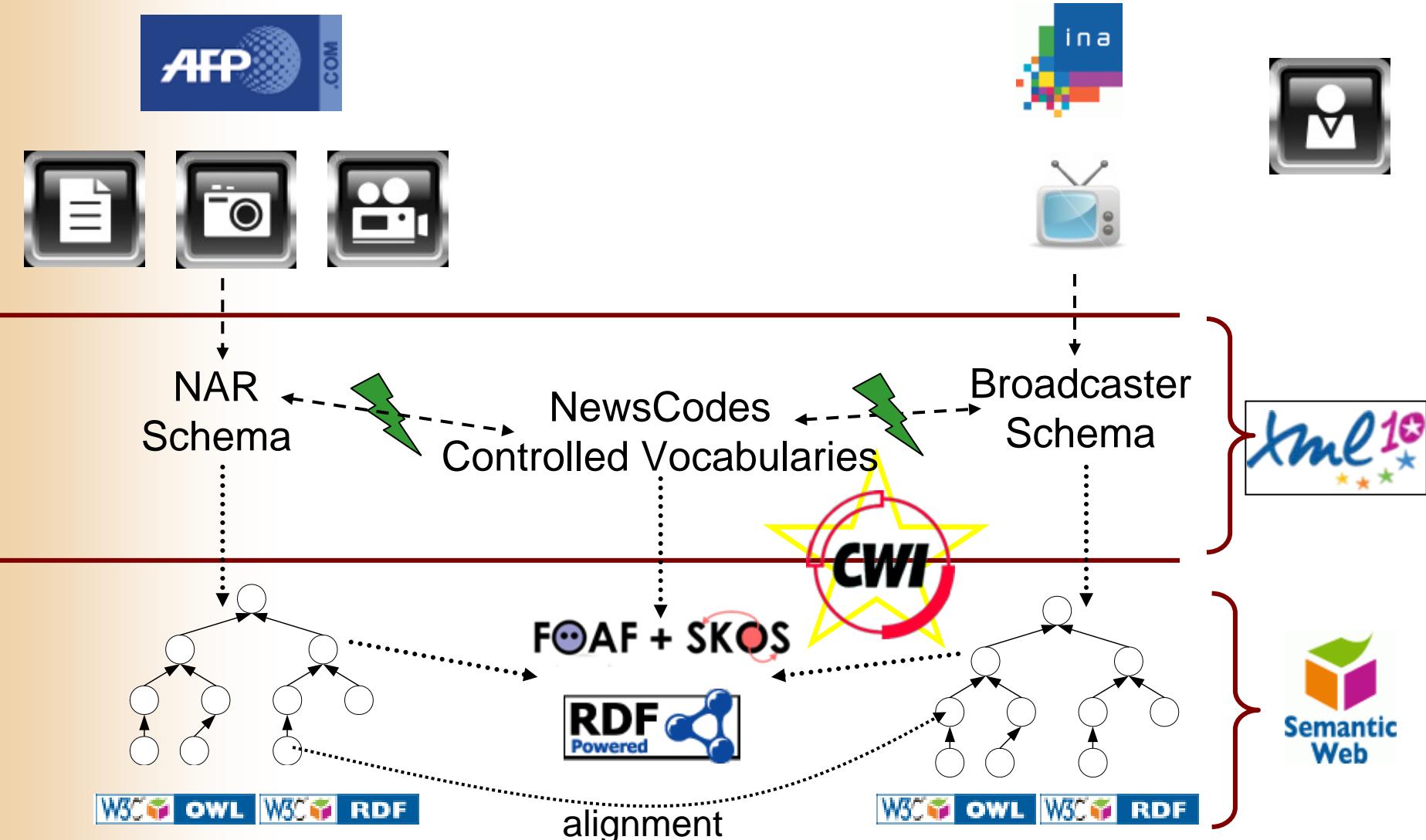
NAR Schema ← - - - - - - - - - → Broadcaster Schema

NewsCodes ← - - - - - - - → Controlled Vocabularies ← - - - → User Vocabulary

Metadata Conversion

Semantic Processing
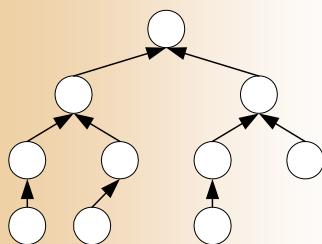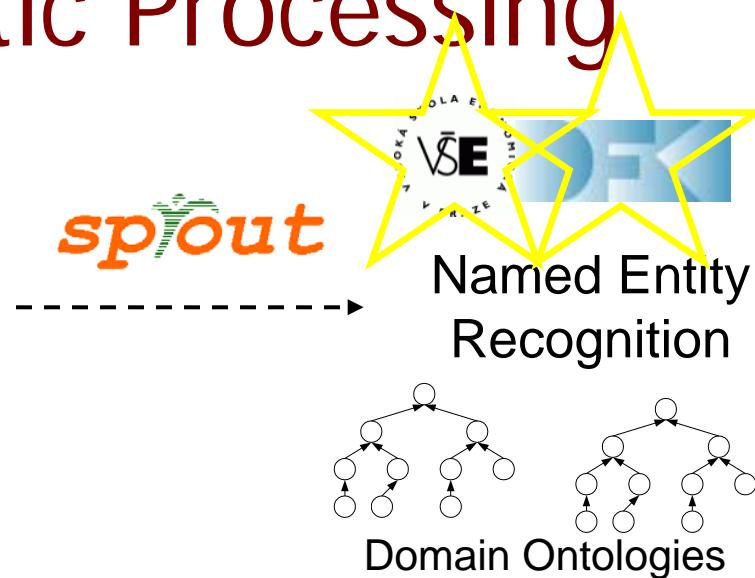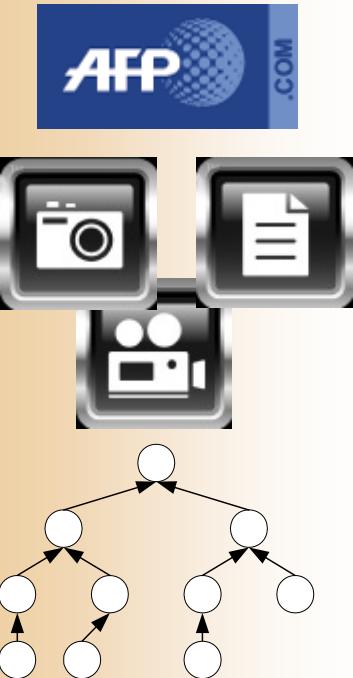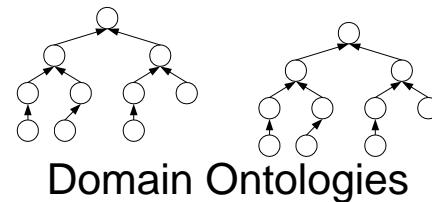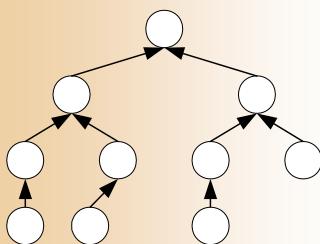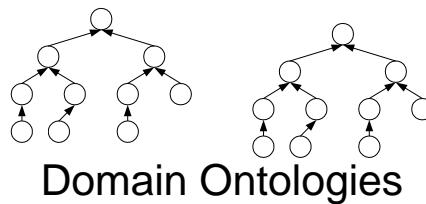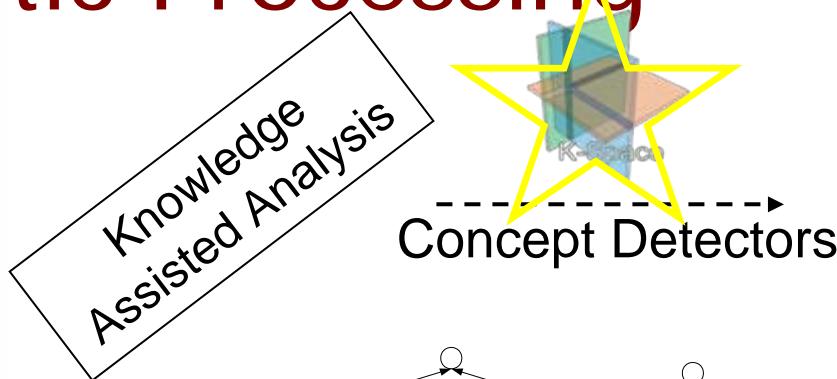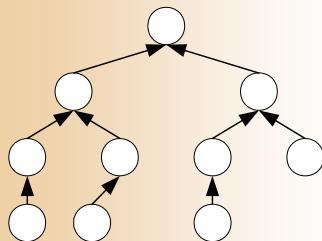
Named Entity
Recognition

Domain Ontologies

NAR Ontology
NewsCodes
Thesaurus

Semantic Processing

Knowledge Assisted Analysis

Concept Detectors

Domain Ontologies

NAR Ontology
NewsCodes
Thesaurus

Semantic Processing

Named Entity Recognition

Linked Data

Domain Ontologies

NAR Ontology
NewsCodes
Thesaurus

Future Work

- Integrate the video browser in the interface
  - Metadata conversion and interoperability
  - Address temporal fragments of the video
  - Visualize videoclips in the interface
- Enrich metadata with visual analysis
  - Apply K-Space concept detectors on visual media
  - Provide new dimensions (facets) for browsing the data
    - Ex: distinguish field images vs stadium and street images with a grass detector for the World Cup dataset
- **Evaluation, Evaluation, Evaluation …**

Literature

- Michiel Hildebrand, Jacco van Ossenbruggen and Lynda Hardman: */facet: A Browser for Heterogeneous Semantic Web Repositories*. In [5th International Semantic Web Conference (ISWC'2006)](), pages 272-285, Athens (GA), USA, November 5-9, 2006.

- Raphaël Troncy, Lynda Hardman, Jacco van Ossenbruggen and Michael Hausenblas: *[Identifying Spatial and Temporal Media Fragments on the Web]()*. In [W3C Video on the Web Workshop](), San Jose (California) and Brussels (Belgium), December 2007.

- W3C Video on the Web Activity, April 2008 [http://www.w3.org/2008/01/video-activity]().

Thanks for your attention

http://www.cwi.nl/~media/www08/