

Multimedia for the Semantic Web

Lynda Hardman*, Jacco van Ossenbruggen, Frank Nack

Multimedia and Human-Computer Interaction

CWI, Amsterdam

** also Technical University, Eindhoven*

- The Multimedia and Human-Computer Interaction group at CWI
- Multimedia as marked up documents and semantic markup
- Media specifics of semantic annotations
- Applications areas of annotated multimedia



Research History of CWI's MM and HCI group

Research on distributed multimedia and the Web for the last 10 years

Different aspects of problem:

- playback environment (multimedia threads)
- distributed sources (requiring synchronization)
- underlying document model (declarative document description, AHM)
- authoring environment (non-programmers can use *GRiNS*)

W3C involvement

- Steven Pemberton (chair XHTML, co-editor CSS1),
- Lloyd Rutledge (co-editor SMIL 1&2, member XLink),
- Jacco van Ossenbruggen (co-editor SMIL 1&2, author SMIL 1&2 DTDs)

ISO involvement

- Frank Nack (MPEG4, MPEG7)



Talk Objectives

Explain the status of multimedia on the Web

Show how current Web specifications can be used in combining multimedia and semantic annotations

Show where semantics are dependent on media type

Indicate direction of our current research



Where semantics and multimedia meet

Ontology (RDF Schema/OIL) a painter is a person who makes paintings	MM specific primitives Co-ordinate space?
Instances (RDF) Rembrandt is a painter	MM specific ontologies scene, sequence, frame defn.
Annotation (XPointer) this video is about Rembrandt	MM annotations this video has N frames
XML data	Multimedia data MPEG2

Marked-up multimedia documents

Ontology (RDF Schema/OIL) a painter is a person who makes paintings	MM specific primitives Co-ordinate space?
Instances (RDF) Rembrandt is a painter	MM specific ontologies scene, sequence, frame defn.
Annotation (XPointer) this video is about Rembrandt	MM annotations this video has N frames
XML data 	Multimedia data MPEG2

Web-based Multimedia Markup—**SMIL**

Synchronized Multimedia Integration Language



SMIL is about timing...



Web-based Multimedia Markup—**SMIL**

Synchronized

Multimedia

Integration

Language

SMIL is about timing...



not just graphics...



Web-based Multimedia Markup—**SMIL**

Synchronized

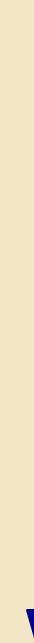
Multimedia

Integration

Language

SMIL is about timing...

not just graphics...



combining Web resources...

Web-based Multimedia Markup—**SMIL**

Synchronized

Multimedia

Integration

Language

SMIL is about timing...

not just graphics...

combining Web resources...

in an XML syntax

A SMIL Document

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE smil PUBLIC "-//W3C//DTD SMIL 1.0//EN"
           "http://www.w3.org/AudioVideo/Group/SMIL10.dtd">
<smil>
  <head>
    <meta name="sync" content="soft"/>
    <layout>
      <root-layout id="SMIL-" width="492" height="810"/>
      <region id="address-region" width="50%" height="8%"/>
      <region id="image-region" top="8%" height="91%"/>
    </layout>
  </head>
  <body>
    <seq>
      <par>
        <text type="text/plain" region="address-region"
              src="Herengracht284.txt" dur="2s"/>
        
      </par>
      <par>
        <text type="text/plain" region="address-region"
              src="Herengracht539.txt"/>
        
      </par>
    </seq>
  </body>
</smil>
```

SMIL in 3 mins

Content

(part of) media item

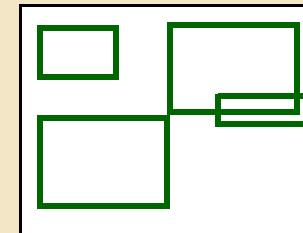
Alternative content
bandwidth
task
user characteristics

Semantic
annotations
meta-data

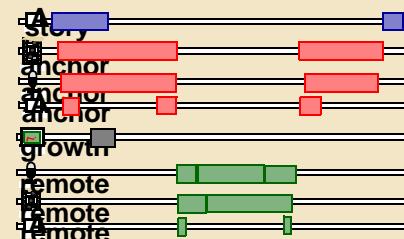
Links
source and destination



Spatial layout



Temporal layout



Basic Semantic Annotations in SMIL

meta element defines properties of a document

```
<meta name="title" content="Web News, 15th June 1998" />  
<meta name="base" content="http://www.cwi.nl/SMIL/webnews/" />
```

- The list of properties (values of **name** attribute) are open-ended.

Attributes on **region** elements

title

Attributes on **par**, **seq** and **media object** elements

abstract, **author**, **copyright**, **title**

Attributes on **media object** elements

alt (contains alternative text)

longdesc (supplement to alt, reference to longer descriptions)

Mixing RDF and SMIL

Ontology (RDF Schema/OIL) a painter is a person who makes paintings	MM specific primitives Co-ordinate space?
Instances (RDF) Rembrandt is a painter	MM specific ontologies scene, sequence, frame defn.
Annotation (XPointer) this video is about Rembrandt	MM annotations this video has N frames
XML data 	Multimedia data MPEG2

Advanced Multimedia Annotations in SMIL

Attaching annotations to different levels in document

- document itself
- elements within document
- media items used in a document

Using the SMIL **metadata** element to embed annotations expressed in:

- RDF
- RDF schema
- OIL
- ... (application dependent ontologies)

Example



Using an existing RDF ontology

For details see <http://www.cwi.nl/~media/semantics/>

```
<?xml version="1.0"?>
<!-- taken from http://www.ics.forth.gr/proj/isst/RDF/RQL/rql.html -->

<rdf:RDF xml:lang="en" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
           xmlns:rdfs="http://www.w3.org/TR/2000/CR-rdf-schema-20000327#"
           xmlns="">

    <rdfs:Class rdf:ID="Artist"/>
    <rdfs:Class rdf:ID="Artifact"/>
    <rdfs:Class rdf:ID="Museum"/>
    <rdfs:Class rdf:ID="Sculptor">
        <rdfs:subClassOf rdf:resource="#Artist"/>
    </rdfs:Class>
    <rdfs:Class rdf:ID="Painter">
        <rdfs:subClassOf rdf:resource="#Artist"/>
    </rdfs:Class>
    <rdfs:Class rdf:ID="Sculpture">
        <rdfs:subClassOf rdf:resource="#Artifact"/>
    </rdfs:Class>
    <rdfs:Class rdf:ID="Painting">
        <rdfs:subClassOf rdf:resource="#Artifact"/>
    </rdfs:Class>

    ...
</rdf:RDF>
```

Embedding RDF in SMIL - I

```
<smil xmlns="http://www.w3.org/2000/SMIL20/CR">
  <head>
    <meta name="generator" content="CWI/Cuypers 1.0" />
    <metadata>
      <rdf:RDF xml:lang="en"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:oil="http://www.ontoknowledge.org/oil/rdf-schema/2000/11/10-oil-standard"
        xmlns:museum="http://ics.forth.gr/.../museum.rdf"
        xmlns:token="http://www.token2000.nl/ontologies/additions" >

        <rdf:Property rdf:about="http://www.token2000.nl/ontologies/additions#painted-by">
          <oil:inverseRelationOf rdf:resource="http://ics.forth.gr/.../museum.rdf#paints"/>
        </rdf:Property>

        <museum:Museum rdf:ID="Rijksmuseum" />

        <museum:Painter rdf:ID="Rembrandt">
          <museum:fname>Rembrandt</museum:fname>
          <museum:lname>Harmenszoon van Rijn</museum:lname>
        </museum:Painter>

        <museum:Painting rdf:about="#apostlePaul">
          <museum:exhibited rdf:resource="#Rijksmuseum" />
          <museum:technique>chiaroscuro</museum:technique>
          <token:painted-by rdf:resource="#Rembrandt" />
        </museum:Painting>

      </rdf:RDF>
    </metadata>
```

Embedding RDF in SMIL - II

```
<museum:Painting rdf:about="#apostlePaul">
  <museum:exhibited rdf:resource="#Rijksmuseum" />
  <museum:technique>chiaroscuro</museum:technique>
  <token:painted-by rdf:resource="#Rembrandt" />
</museum:Painting>

</rdf:RDF>
</metadata>
...
</head>
<body>
  <par>
    <text region="title" src="...query to multimedia database..." />
    <text region="descr" src="..." />
    <seq>
      <par dur="10"> ... 1st painting+title ... </par>
      <par dur="10"> ... 2nd painting+title ... </par>
      <par dur="10"> ... 3rd painting+title ... </par>
      <par dur="10"> ... 4th painting+title ... </par>
      <par dur="10" id="apostlePaul">
        
        <text region="ptitle" src="..." />
      </par>
    </seq>
  </par>
</body>
</smil>
```

Schemas for Multimedia

Ontology (RDF Schema/OIL)
a painter is a person who makes paintings

Instances (RDF)
Rembrandt is a painter

Annotation (XPointer)
this video is about Rembrandt

XML data

MM specific primitives
Co-ordinate space?

MM specific ontologies
scene, sequence, frame defn.

MM annotations
this video has N frames

Multimedia data
MPEG2

Schemas for Multimedia

RDF Schema allows the definition of application specific schemas

- is it sufficiently powerful for describing multimedia specific schemas?
- or do we need multimedia specific modeling primitives in RDFS (e.g. time)?

RDF/XML Schema provide generic data types

- e.g. strings, integers, dates
- We need to specify data types geared to multimedia content
 - sample rates and sizes, colour spaces (rgb, hsv, cmyk), loudness levels (dB)
- Media-specific data types, e.g. video
 - scene, sequence, shot, frame

In either case, for interoperability multimedia needs a common schema as a basis for application specific schemas

- “Dublin Core” for multimedia

This is not easy...

- five MPEG committee members results in 7 different schemas

Annotating Multimedia

Ontology (RDF Schema/OIL)
a painter is a person who makes paintings

Instances (RDF)
Rembrandt is a painter

Annotation (XPointer)

this video is about Rembrandt

XML data

MM specific primitives
Co-ordinate space?

MM specific ontologies
scene, sequence, frame defn.

MM annotations

this video has N frames

Multimedia data

MPEG2

Assigning semantic annotations to media

To what do we need to attach the annotations?

- specific player in football match
- audio fragment of gun shot
- third character out of shot
- specified bounding box in frame smpte=13:21:33:20
- video currently playing in active window
- the last viewed video

Requires media-dependent descriptions of media fragments

- “XPointer” for media types other than text/xml

The good old hypermedia anchoring questions revisited...



Combining distributed ontology fragments

On the Web everyone, including multimedia, needs...

more than the traditional KR approach:

- ontology is assumed to be complete, consistent and authoritative

but on the semantic web:

- ontologies are potentially fragmented, inconsistent and unreliable

Want to combine ontology fragments within a single media item description

Collecting annotations

Pre production

- storyboard
- script

During production

- camera
 - lens (zoom), focus, shutter, colour histogram analysis, tracker (pan, angle)
 - simple high-level semantics (quality)

Post production

- manually add annotations
- edit lists (path through material)
- automatic feature extraction

A4SM (GMD-IPSI project) developed MPEG7 camera

Application Areas

Information retrieval and filtering

- pattern matching at data level - string or histogram
- high-level search on concepts

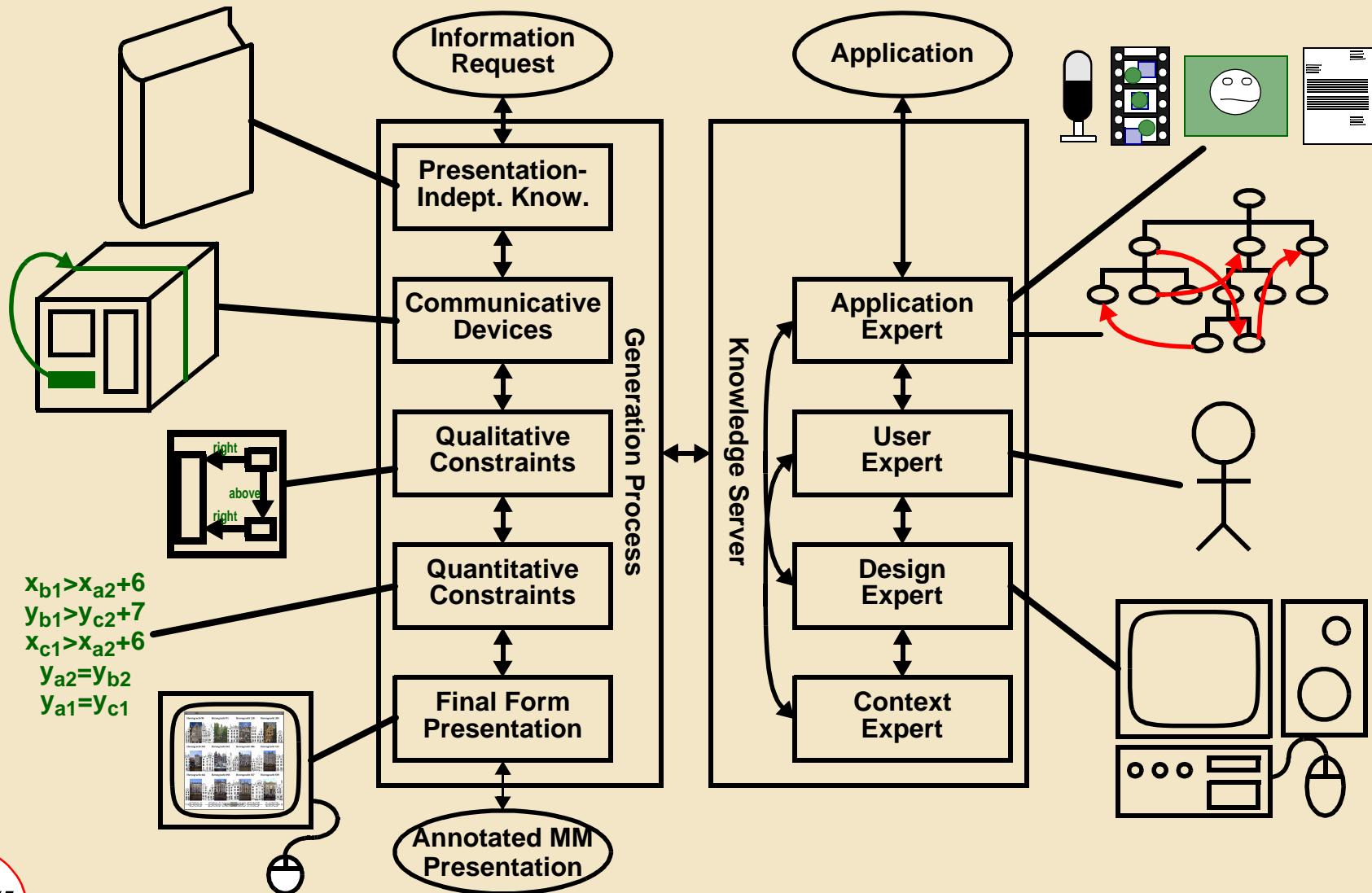
Surveillance

- low-level real-time filtering
looking for low-level patterns which may indicate intrusion

Automated presentation generation

Semantic markup generation

Overall Generation Process Architecture



Presentation Independent Knowledge

Ontologies to describe the underlying application domain

- e.g. from museum ontology
 - painter, painted-by, chiaroscuro

Ontologies to describe the means of communication (e.g. narrative, rhetoric)

- e.g. from Rhetorical Structure Theory (RST, Mann & Thompson '89), include
 - elaboration, motivation, sequence

Information from both types is taken into account when generating
communicative devices



Communicative devices

A **communicative device** is a frequently-used presentation technique

- e.g. a hypermedia design pattern

A rhetorical relation can be presented using different communicative devices

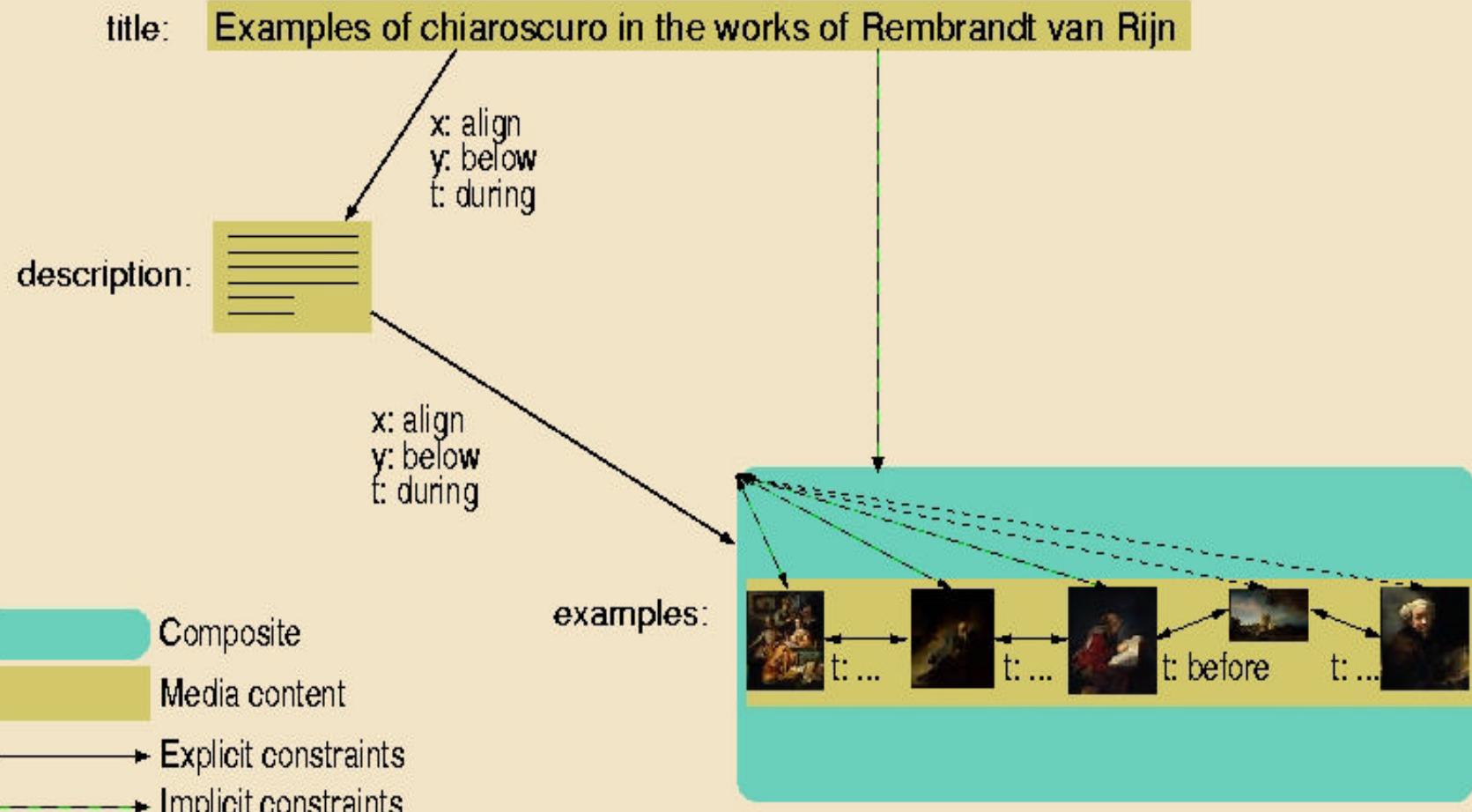
- For example, “sequence” can be presented using a bookshelf communicative device

Bookshelf communicates the order of the elements of a sequence using the most appropriate dimension(s)

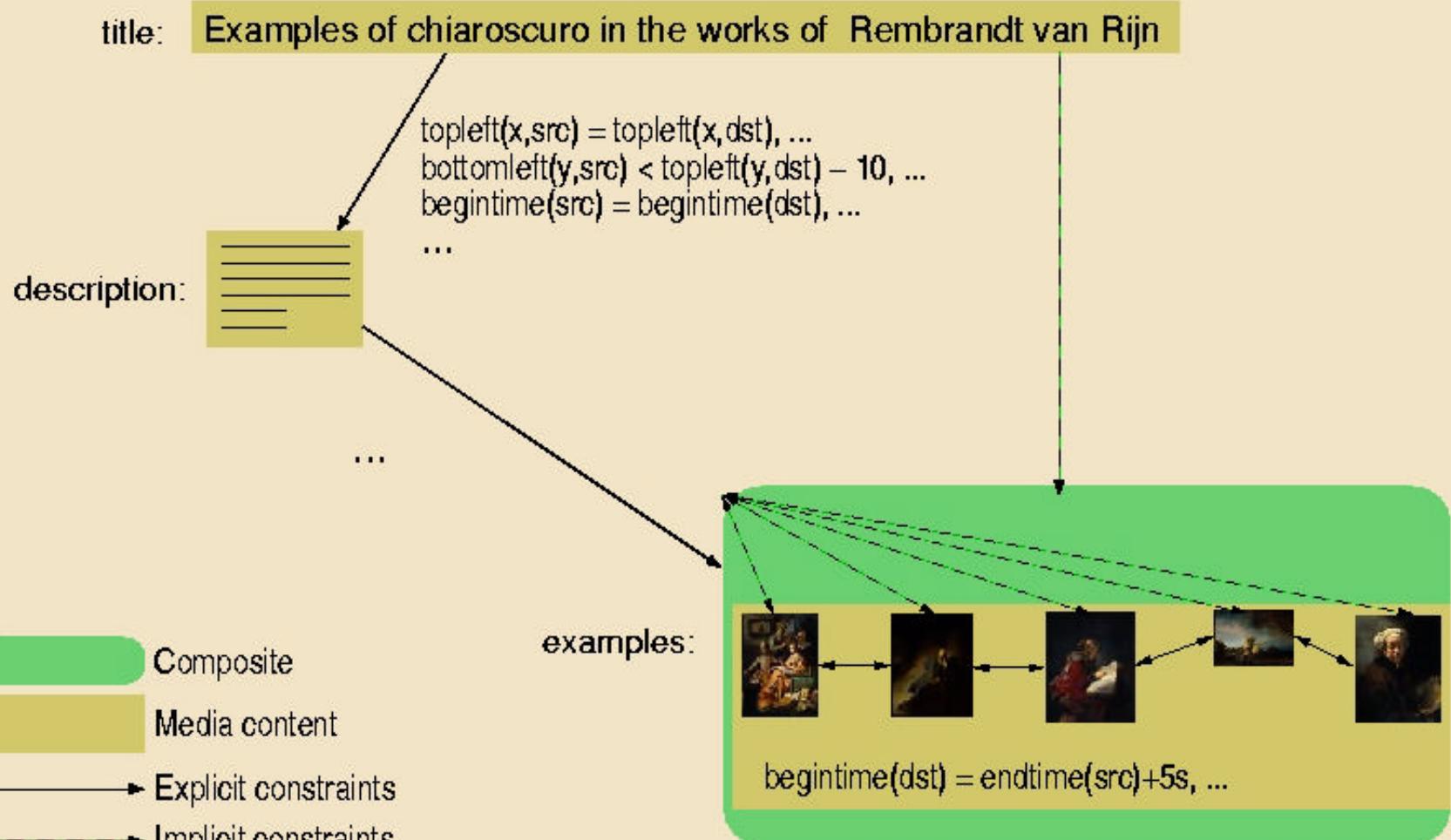
- spatial (order on page)
- temporal (one after the other)
- links (previous/next buttons)

Communicative devices are specified in terms of qualitative constraints

Qualitative constraints



Quantitative constraints



Final-form presentation



```
<museum:Painting rdf:about="#apostlePaul">
  <museum:exhibited rdf:resource="#Rijksmuseum" />
  <museum:technique>chiaroscuro</museum:technique>
  <token:painted-by rdf:resource="#Rembrandt" />
</museum:Painting>

</rdf:RDF>
</metadata>
...
</head>
<body>
<par>
  <text region="title" src="...query to multimedia da...
  <text region="descr" src="..."/>
  <seq>
    <par dur="10"> ... 1st painting+title ... </par>
    <par dur="10"> ... 2nd painting+title ... </par>
    <par dur="10"> ... 3rd painting+title ... </par>
    <par dur="10"> ... 4th painting+title ... </par>
    <par dur="10" id="apostlePaul">
      
      <text region="ptitle" src="..."/>
    </par>
  </seq>
</par>
</body>
</smil>
```

Conclusions

Multimedia puts specific demands on the Semantic Web

Semantic annotations for multimedia presentation generation

- collecting annotations as part of the input to the generation process
- generating better annotations as part of the output

Some of the hard problems are:

- how to link down into the media-specific level
- how to come up with agreed-upon media-specific descriptions
- how to combine ontology fragments
- do schema languages need built-in multimedia modelling primitives
- how can annotation collection be made part of the production process