

# Multimedia on the Semantic Web

Jacco van Ossenbruggen, Lynda Hardman, Frank Nack

*Multimedia and Human-Computer Interaction*  
*CWI, Amsterdam*

- Short history of the Web in three generations (see thesis for long history)
- Multimedia as marked up documents including semantic markup
- Media specifics of semantic annotations

## Talk Objectives

### Crash course Semantic Web

- RDF, RDF Schema, PICS, P3P, DAML+OIL, ...

### What can we do now:

- Multimedia on the Web
- Multimedia and semantic annotations

### What we want but can't do yet

- Show where the Semantic Web needs media-specific functionality

### Indicate direction of our current research

- role of semantics in hypermedia presentation generation

## The Web: First Generation

### **Problem: hard to find and get access to information on the internet**

- Different platforms (Unix, PC, Mac)
- Different wordprocessing and typesetting software
- Different proprietary file formats and encodings (compression)
- Different naming rules for file pathnames
- Command line syntax of transfer applications (FTP)
- ...

### **Solution: The browser as the single interface to all information**

- Platform independent
- HTML standardized the document format
- URIs standardized the naming rules
- Hypertext links replaced command line syntax
- ...

## The Web: Second Generation

### **Problem: Amount of manual effort related to Web authoring**

- Expensive and time consuming (not suited for volatile content)
- Inflexible and hard to maintain
- Limited (re)use of database content
- ...

### **Solution: Automate production of Web pages**

- Cheap and fast (even on the fly generation)
- Flexible
- Make all database content accessible over the Web
- ...

## The Web: Third Generation

### **Problem: Web content only interpretable by humans**

- hardly interpretable by machines
- webbots become “screen scrapers”
- amount of “semantics” extracted is too limited (and too much work)
- ...

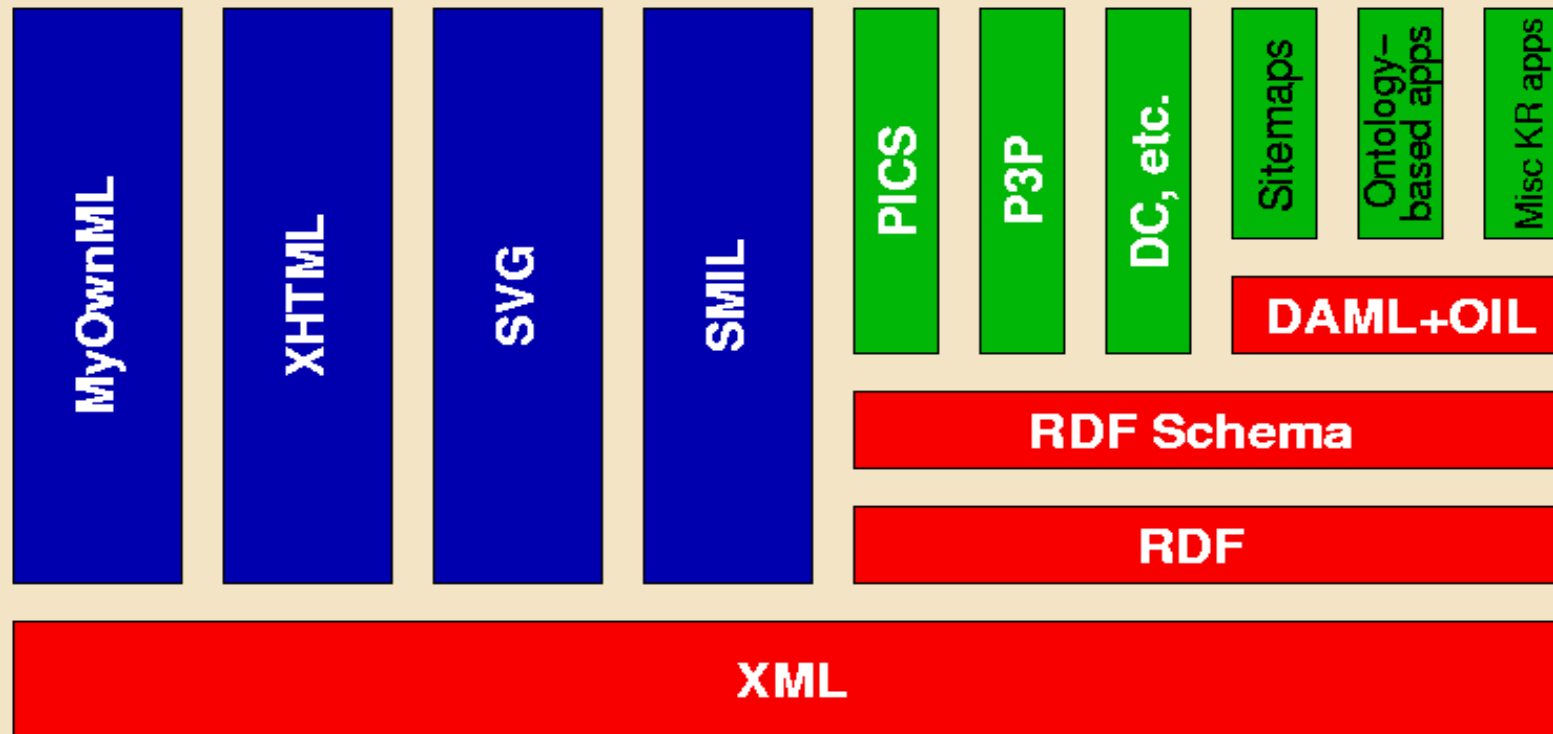
### **Solution: make semantics of Web content explicit**

- add machine-readable annotations to Web content
- provide (formal) definitions for the meaning of these annotations
- provide infrastructure to make this scalable and interoperable over the Web
- ...

# Document vs. Knowledge Representation on the Web

## document representation

## knowledge representation



# Overview Semantic Web

## XML (Rec)

- common syntactical layer enables reuse of XML parsers, XPath, XPointer, XSLT, CSS, ...

## RDF (Rec)

- common data model for simple statements about Web resources
- everything is a triple: (subject, predicate, object) or (resource, property, value)
- simple collections (bags, sequences and alternatives)
- reification (statements about statements)

## RDF Schema (CR)

- schema language for defining RDF vocabularies
- gives class/subclass hierarchy and properties with domain/range restrictions

## DAML+OIL

### DARPA project building on

- European results, and is also a topic in the just started
- W3C Semantic Web activity

### Adds elements common in frame systems and description logics

- local properties
- cardinality constraints
- logical class expressions
- ...

### Formal semantics and fast implementations

- except reification



## So where is the multimedia?

### Still trying to get past the first generation Web

- proprietary data formats (even if SMIL is used)
- manually authored

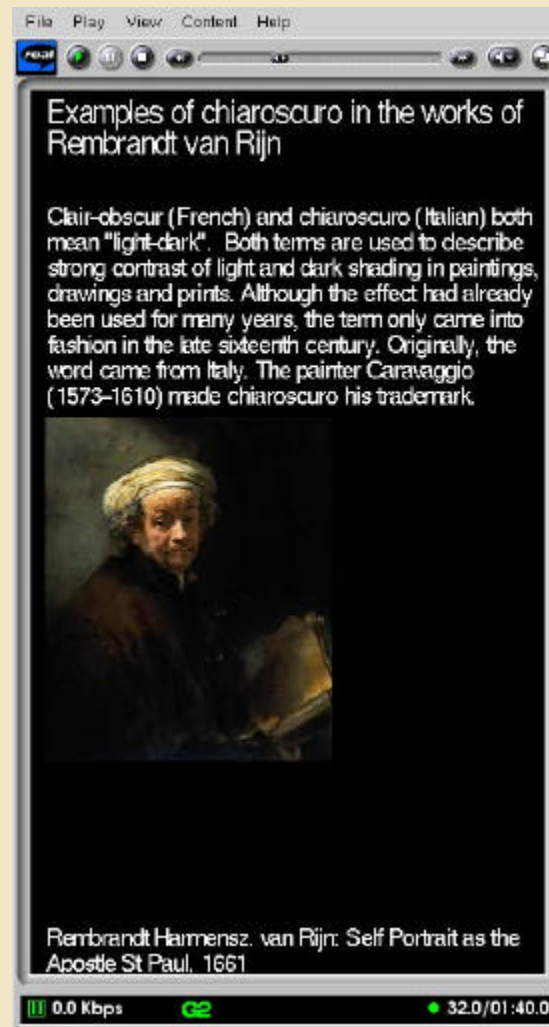
### At CWI, we are doing second generation Web...

- automatic generation of multimedia presentations
- no templates or functional transformation rules but ...
- ... rules that use backtracking and constraint solving
- see WWW10 paper for details

### ... and heading towards third generation Web

- generation of annotated multimedia
- reuse of semantics needed for generation process
- based on current state of the art
- need for multimedia specific semantics

## 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 (1)

```

<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-oilstandard"
        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>
  </head>
</smil>

```

## Embedding RDF in SMIL (2)



```

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

</rdf:RDF>
</metadata>
...
</head>
<body>
  <par>
    <text region="title" src="...query to multimedia data" />
    <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>

```

## Where multimedia and semantics meet

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

# Mixing RDF and SMIL

## 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

SMIL

## MM specific primitives

Co-ordinate space?

## MM specific ontologies

scene, sequence, frame defn.

## MM annotations

this video has N frames

Multimedia data

MPEG2

# Annotating Multimedia

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



## 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 and linking questions revisited...

- to embed or not to embed
- annotation server infrastructure
- ...

# 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

## Combining distributed ontology fragments

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

**More than the traditional KR approach:**

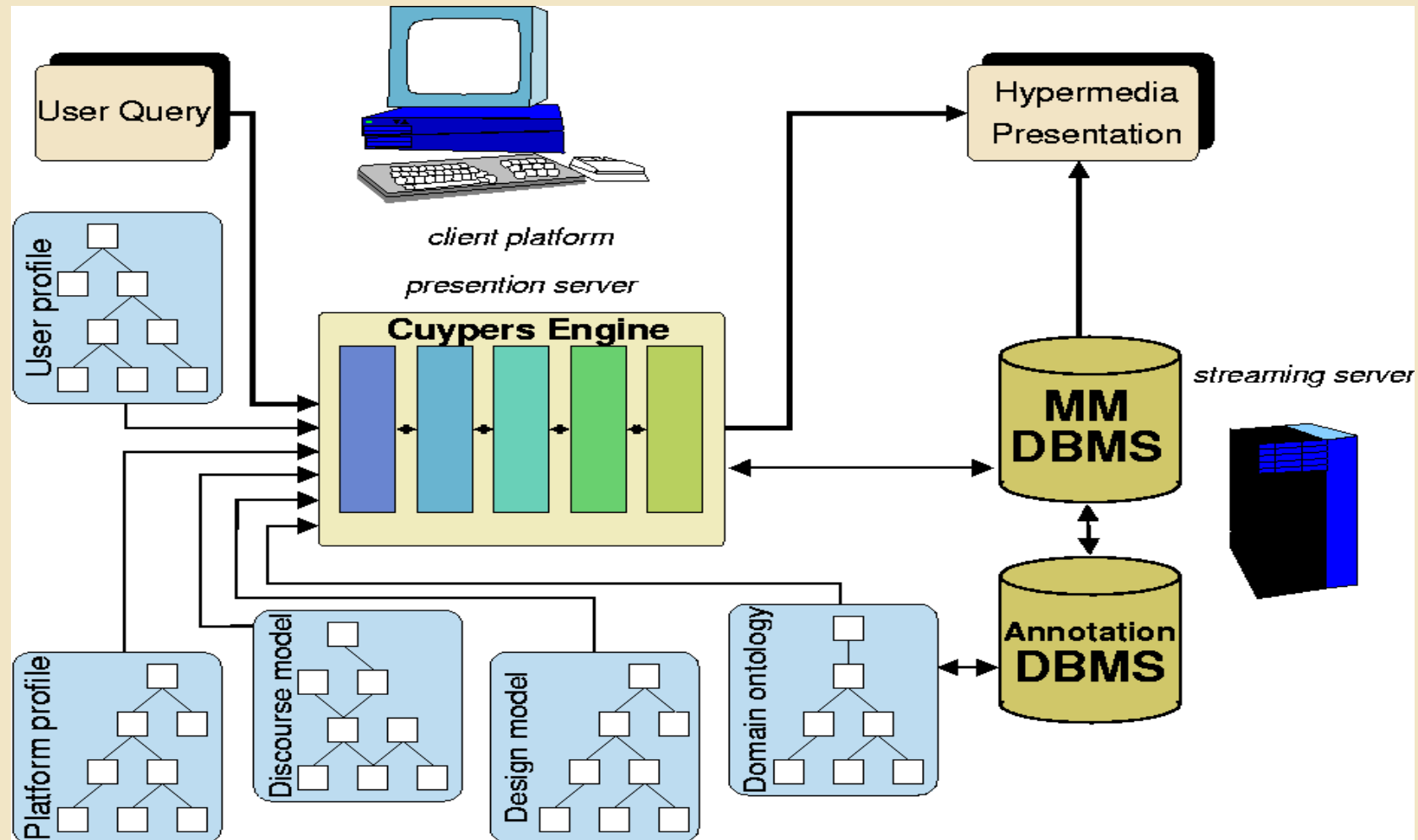
- ontology is assumed to be complete, consistent and authoritative

**But on the semantic web:**

- ontologies are potentially fragmented, inconsistent and unreliable

**Need to combine ontology fragments within a single media item description**

# Semantics for hypermedia presentation generation



## 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

## Pointers

### Specifications:

- [www.w3.org](http://www.w3.org)
- [www.daml.org](http://www.daml.org)

### CWI's multimedia group

- WWW10 paper:
  - [www.cwi.nl/~media/publications/www10/](http://www.cwi.nl/~media/publications/www10/)
- SMIL example:
  - [www.cwi.nl/~media/semantics/](http://www.cwi.nl/~media/semantics/)
- My thesis
  - [www.cwi.nl/~jrvosse/thesis/](http://www.cwi.nl/~jrvosse/thesis/)