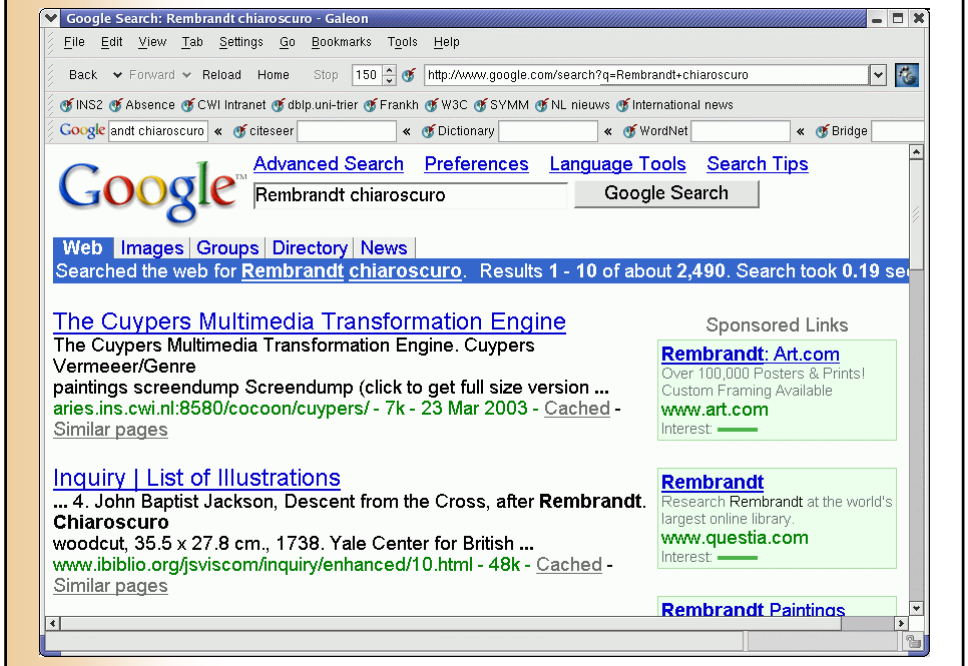
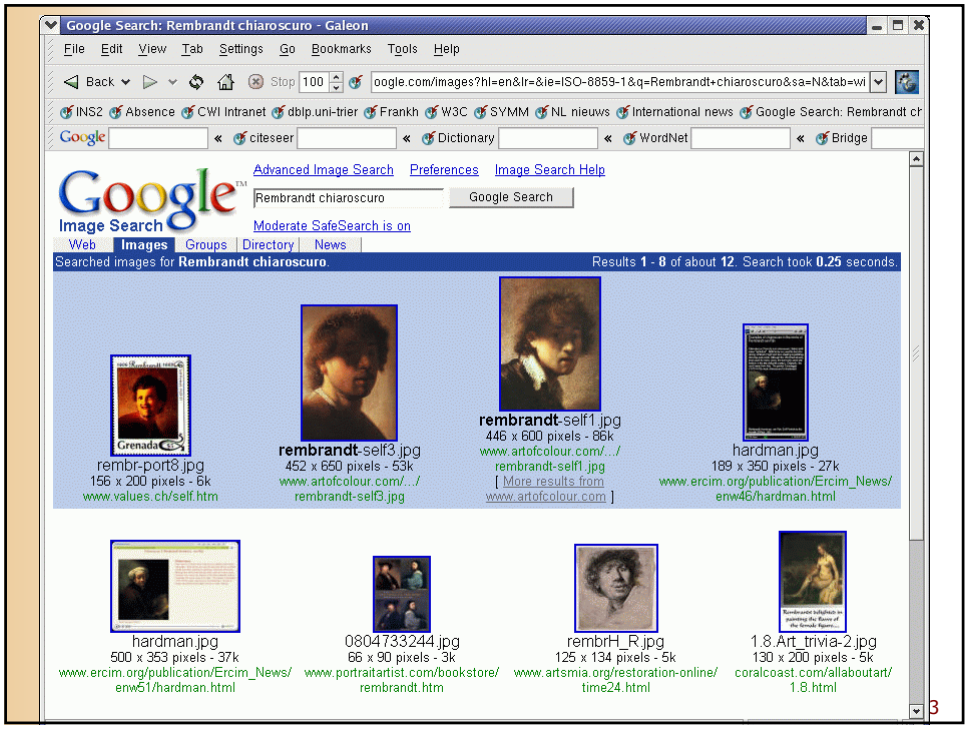
Providing Flexible Interfaces to Annotated Multimedia Repositories

Lynda Hardman

CWI, Semantic Media Interfaces
TU/e, Multimedia and Internet Technology



2


Google Search: Rembrandt chiaroscuro - Galeon
File Edit View Tab Settings Go Bookmarks Tools Help
Back Stop 100
oogle.com/images?hl=en&lr=&ie=ISO-8859-1&q=Rembrandt+chiaroscuro&sa=N&tab=wi
INS2 Absence CWI Intranet dblp.uni-trier Frankh W3C SYMM NL nieuws International news Google Search: Rembrandt ch
Google citeseer Dictionary WordNet Bridge
Google Image Search
Advanced Image Search Preferences Image Search Help
Rembrandt chiaroscuro
Google Search
Moderate SafeSearch is on
Web Images Groups Directory News
Searched images for Rembrandt chiaroscuro.
Results 1 - 8 of about 12. Search took 0.25 seconds.
rembr-port8.jpg
156 x 200 pixels - 6k
www.values.ch/self.htm
rembrandt-self3.jpg
452 x 650 pixels - 53k
www.artofcolour.com/.../rembrandt-self3.jpg
rembrandt-self1.jpg
446 x 600 pixels - 86k
www.artofcolour.com/.../rembrandt-self1.jpg
[More results from www.artofcolour.com]
hardman.jpg
189 x 350 pixels - 27k
www.ercim.org/publication/Ercim_News/enw46/hardman.html
hardman.jpg
500 x 353 pixels - 37k
www.ercim.org/publication/Ercim_News/enw51/hardman.html
0804733244.jpg
66 x 90 pixels - 3k
www.portraitartist.com/bookstore/rembrandt.htm
rembrH_R.jpg
125 x 134 pixels - 5k
www.artsmia.org/restoration-online/time24.html
1.8.Art_trivia-2.jpg
130 x 200 pixels - 5k
coralcoast.com/allaboutart/1.8.html

3


Chiaroscuro & Rembrandt Harmensz. van Rijn
Chiaroscuro
Self Portrait (1661)

4

One size *doesn't* fit all

5

The problem

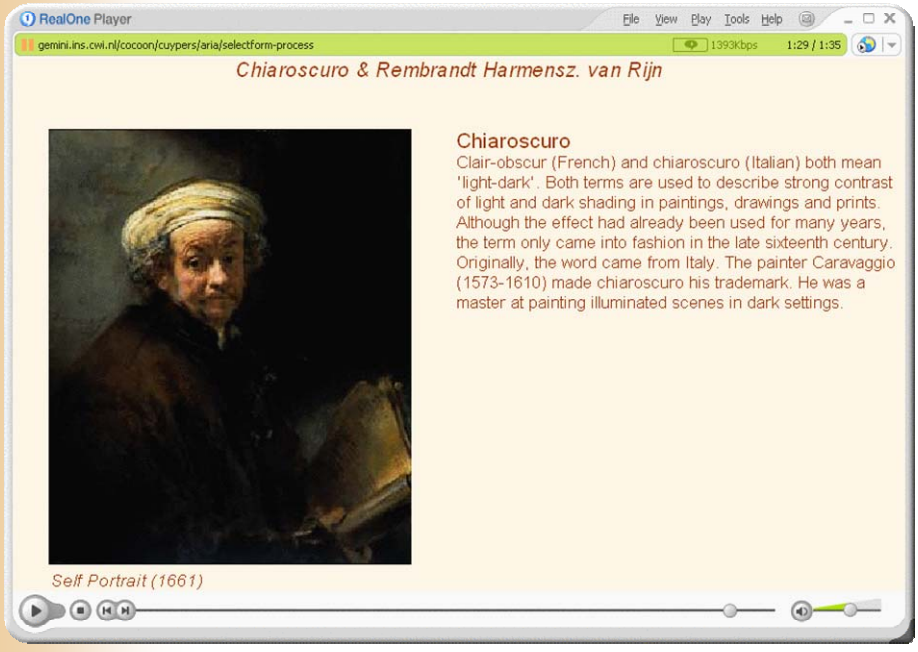
- Too many users need their own information
  - for their level of expertise
  - using appropriate media
  - in an appropriate style
  - displayed on their own device
- Multimedia information design is expensive
- There has to be some automation in the process

6

Overview of topics

- Dependencies in multimedia design
- Multimedia documents on the semantic web
- Formalising annotations for audiovisual documents
- NewsML 2.0: semantics of news media assets
- Vox Populi: creating argument structure with video fragments

7



8

Three ingredients

Content

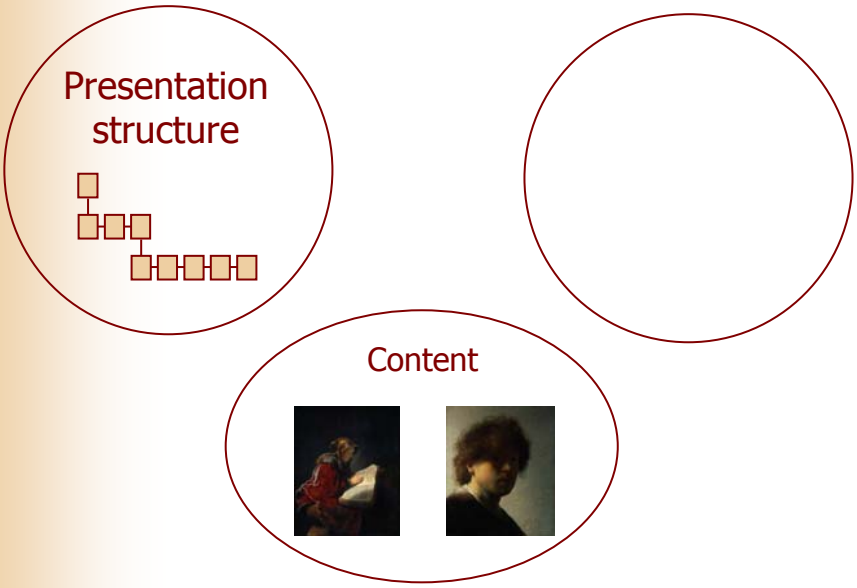

9

Content of example


Clair-obscur (Frans) en chiaroscuro (Italiaans) betekenen 'licht-donker'. Beide termen worden gebruikt om sterke licht-donkercontrasten in schilderijen, tekeningen en prenten aan te duiden. Hoewel het effect al eerder werd toegepast, is de term pas sinds het einde van de 16de eeuw in zwang. De oorsprong van het woord ligt in Italië. De schilder Caravaggio (1573-1610) maakte het chiaroscuro-effect tot zijn handelsmerk. Hij was een meester in het schilderen van donkere taferelen met één felle lichtbundel.

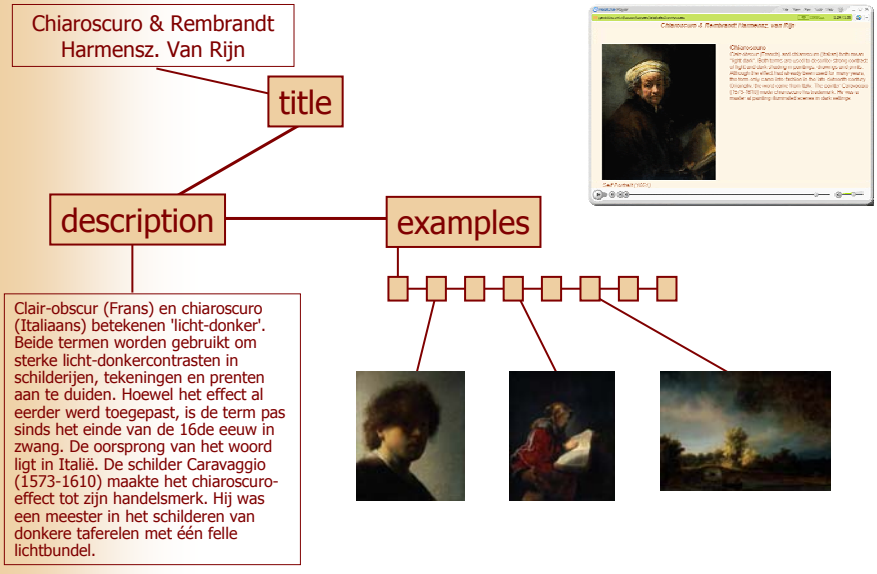

10

Three ingredients



11

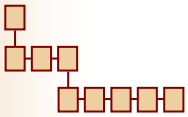
Presentation structure of example



12

Three ingredients

Presentation structure



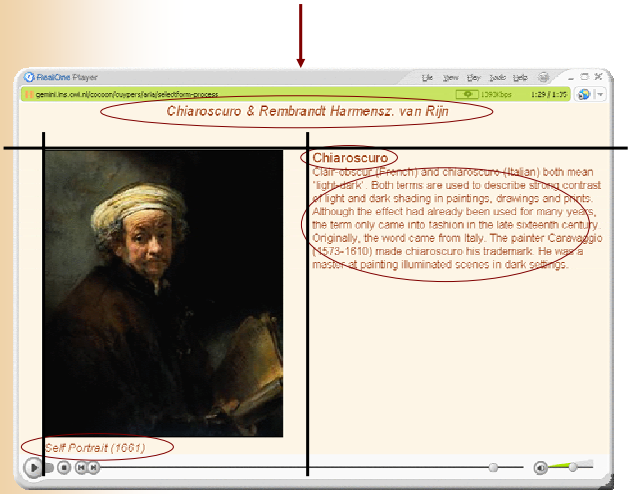
Aesthetics

abcdefghij
abcdefghijklmn
ABCDEFGHIJK
abcdefghijklm
abcdefghijklm

Content

13

Aesthetics of example



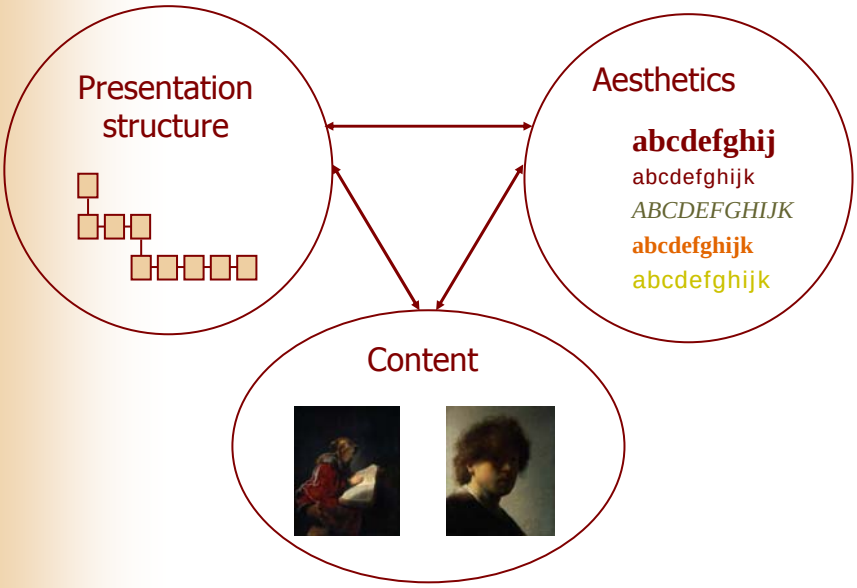
Fonts

Layout

Colours

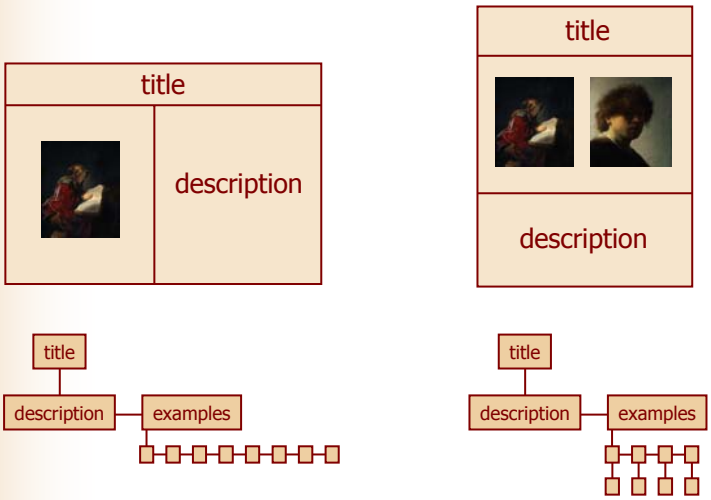
14

Design dependencies



15

Presentation Structure depends on Layout



16

Style Depends on Content

Chiaroscuro & Rembrandt

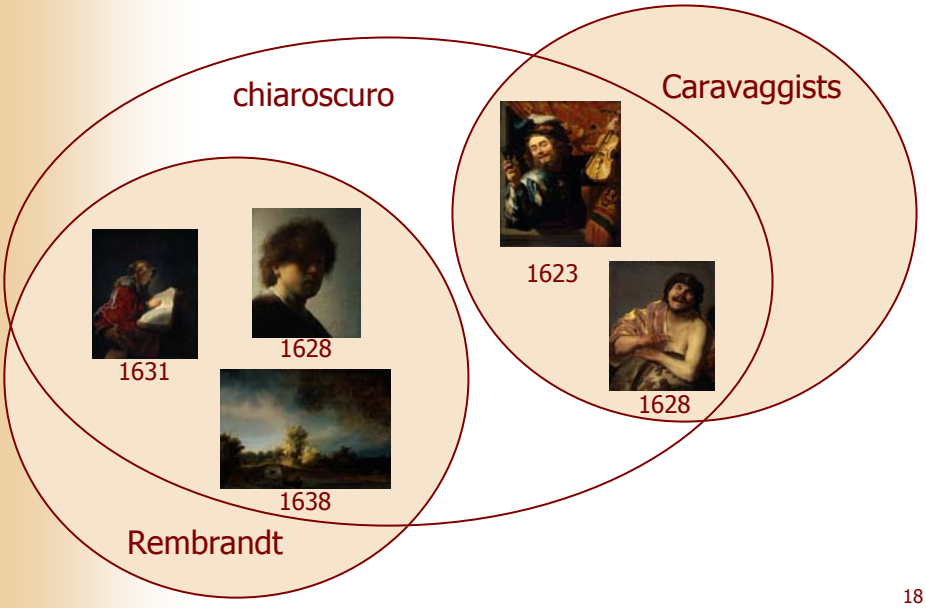
The Stone Bridge (1638)

Chiaroscuro & Rembrandt

The Stone Bridge (1638)

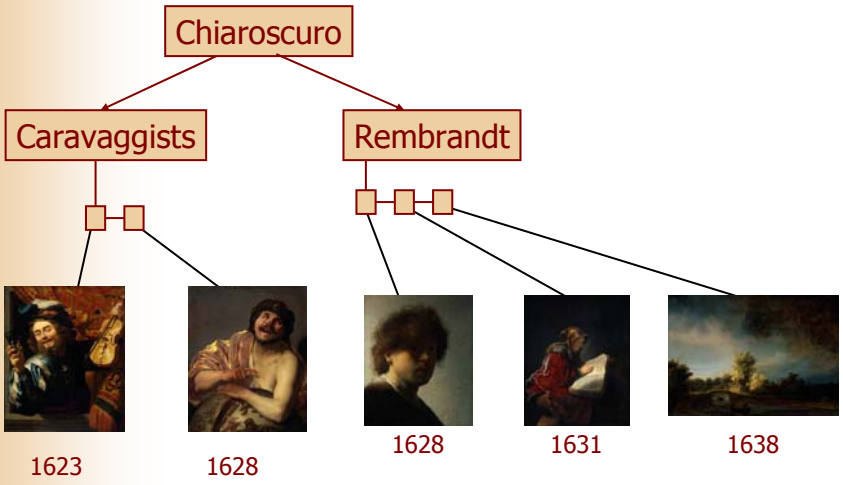
17

Presentation structure depends on content



18

Example Presentation Structure

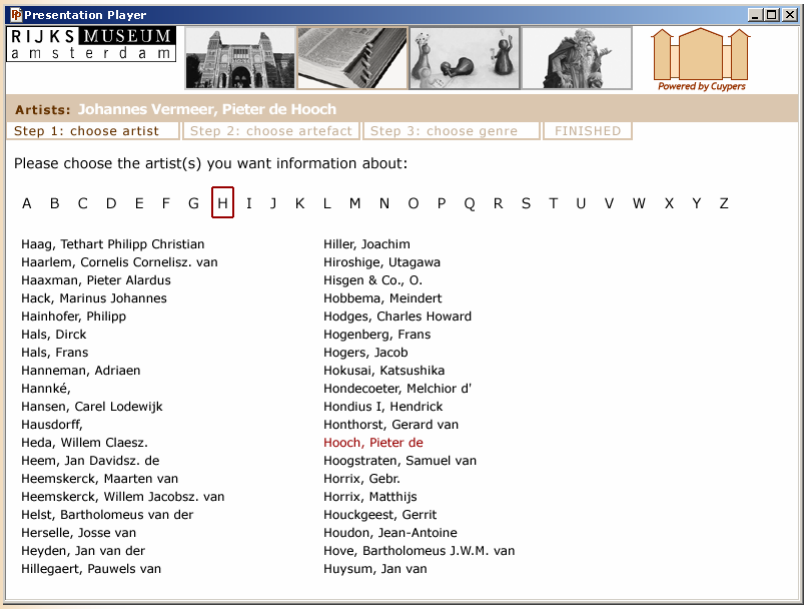


19

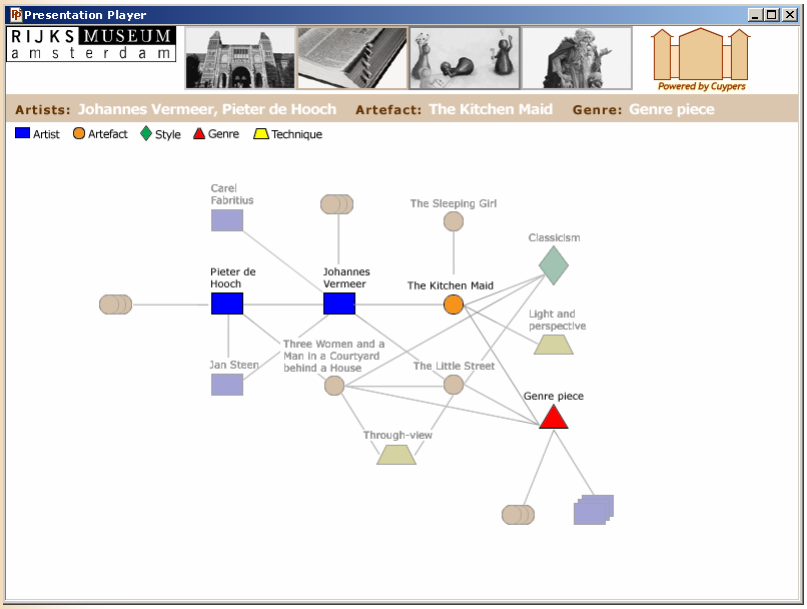
Different presentation styles

- Large amount of information
- High interaction

20


Presentation Player
RIJKS MUSEUM
a m s t e r d a m
Powered by Cuypers
Artists: Johannes Vermeer, Pieter de Hooch
Step 1: choose artist
Step 2: choose artefact
Step 3: choose genre
FINISHED
Please choose the artist(s) you want information about:
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Haag, Tethart Philipp Christian
Haarlem, Cornelis Cornelisz. van
Haaxman, Pieter Alardus
Hack, Marinus Johannes
Hainhofer, Philipp
Hals, Dirck
Hals, Frans
Hanneman, Adriaen
Hannké,
Hansen, Carel Lodewijk
Hausdorff,
Heda, Willem Claesz.
Heem, Jan Davidsz. de
Heemskerck, Maarten van
Heemskerck, Willem Jacobsz. van
Helst, Bartholomeus van der
Herselle, Josse van
Heyden, Jan van der
Hillegaert, Pauwels van
Hiller, Joachim
Hiroshige, Utagawa
Hisgen & Co., O.
Hobbema, Meindert
Hodges, Charles Howard
Hogenberg, Frans
Hogers, Jacob
Hokusai, Katsushika
Hondecoeter, Melchior d'
Hondius I, Hendrick
Honthorst, Gerard van
Hooch, Pieter de
Hoogstraten, Samuel van
Horrix, Gebr.
Horrix, Matthijs
Houckgeest, Gerrit
Houdon, Jean-Antoine
Hove, Bartholomeus J.W.M. van
Huysum, Jan van

21


Presentation Player
RIJKS MUSEUM
a m s t e r d a m
Powered by Cuypers
Artists: Johannes Vermeer, Pieter de Hooch
Artefact: The Kitchen Maid
Genre: Genre piece
Artist
Artefact
Style
Genre
Technique
Carel Fabritius
The Sleeping Girl
Classicism
Pieter de Hooch
Johannes Vermeer
The Kitchen Maid
Light and perspective
Three Women and a Man in a Courtyard behind a House
Jan Steen
The Little Street
Genre piece
Through-view

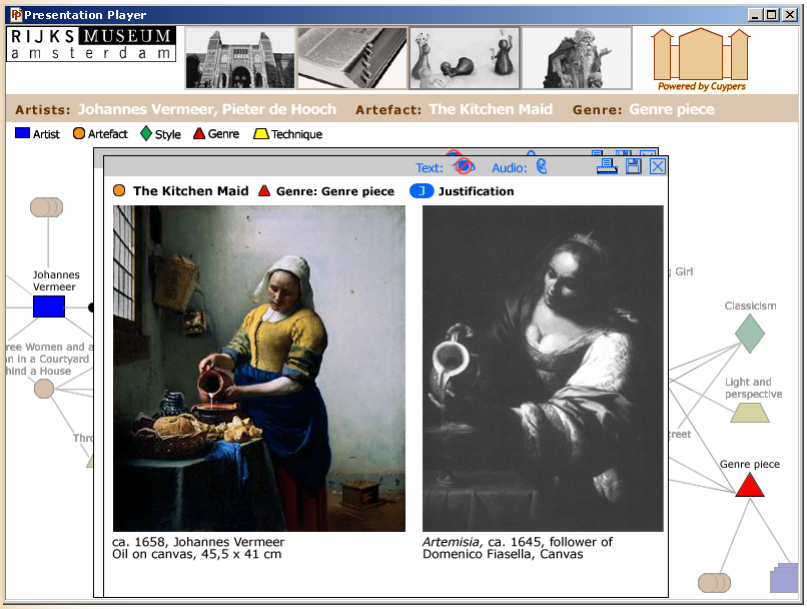
22


Presentation Player
RIJKS MUSEUM
a m s t e r d a m
Powered by Cuypers
Artists: Johannes Vermeer, Pieter de Hooch
Artefact: The Kitchen Maid
Genre: Genre piece
Artist
Artefact
Style
Genre
Technique
1 2 3
Text:
Audio:
The Kitchen Maid
With quiet concentration a woman pours milk into a bowl. With her left hand she supports the can she is pouring from. Around her are various objects: a loaf of bread, a stoneware jug, a basket and a brass bucket. The woman is standing near the window so she can see what she is doing. The light falls on her hands; her silhouette is dark against the white wall. There is a fascinating play of light and shadow in this painting.
Genre
This is one of Johannes Vermeer's genre pieces in which he establishes an intensely intimate atmosphere. Although the artist observes his model from nearby, she continues with her work, totally unperturbed.
E
Ex
J
All
Technique
Vermeer made use of light and perspective to create the intimate atmosphere. All lines of perspective lead to the right hand of the girl, which subtly accentuates the task of pouring milk in which she is completely engrossed. The horizon lies beneath her head, so that the viewer seems to look up
ca. 1658, Johannes Vermeer
Oil on canvas, 45,5 x 41 cm
Johannes Vermeer
Classicism
Light and perspective
Street
Genre piece

23


Presentation Player
RIJKS MUSEUM
a m s t e r d a m
Powered by Cuypers
Artists: Johannes Vermeer, Pieter de Hooch
Artefact: The Kitchen Maid
Genre: Genre piece
Artist
Artefact
Style
Genre
Technique
Text:
Audio:
The Kitchen Maid
ca. 1658, Johannes Vermeer
Oil on canvas, 45,5 x 41 cm
Genre
Technique
Style
Johannes Vermeer
The Sleeping Girl
E
Ex
J
All
Johannes Vermeer
Classicism
Light and perspective
Street
Genre piece

24



25

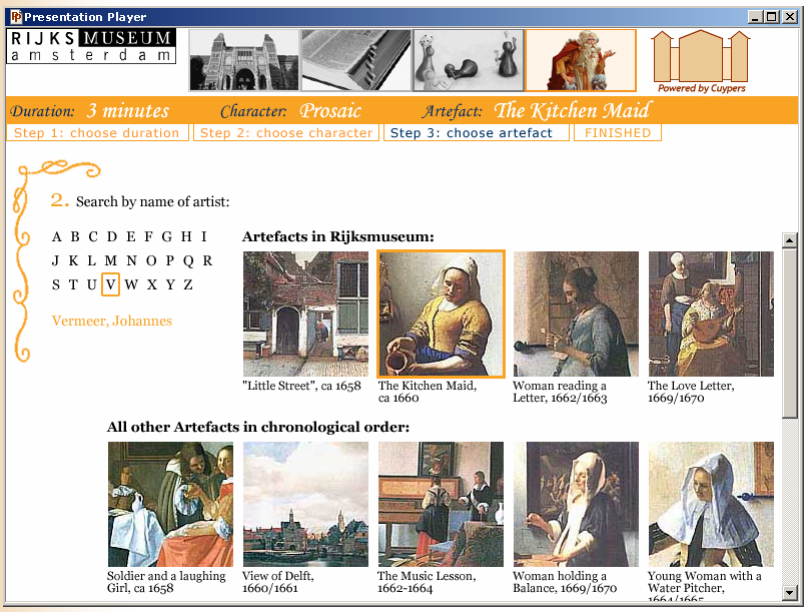
Different presentation styles

- Entertainment rather than information
- Low interaction

26


Presentation Player
RIJKS MUSEUM
a m s t e r d a m
Powered by Cuypers
Duration: 3 minutes
Character: Prosaic
Step 1: choose duration
Step 2: choose character
Step 3: choose artefact
FINISHED
Step 1: How long do you want your story to be? 3 minutes.
Step 2: What is the preferred character of your story?
Prosaic

27


Presentation Player
RIJKS MUSEUM
a m s t e r d a m
Powered by Cuypers
Duration: 3 minutes
Character: Prosaic
Artefact: The Kitchen Maid
Step 1: choose duration
Step 2: choose character
Step 3: choose artefact
FINISHED
2. Search by name of artist:
A B C D E F G H I
J K L M N O P Q R
S T U V W X Y Z
Vermeer, Johannes
Artefacts in Rijksmuseum:
"Little Street", ca 1658
The Kitchen Maid, ca 1660
Woman reading a Letter, 1662/1663
The Love Letter, 1669/1670
All other Artefacts in chronological order:
Soldier and a laughing Girl, ca 1658
View of Delft, 1660/1661
The Music Lesson, 1662-1664
Woman holding a Balance, 1669/1670
Young Woman with a Water Pitcher,

28

Presentation Player

RIJKS MUSEUM amsterdam

Powered by Cuypers

Duration: 3 minutes *Character:* Prosaic *Artefact:* The Kitchen Maid

Text: Audio: Speed:

The Kitchen Maid
by Johannes Vermeer
ca 1660
Oil on canvas
45.5 x 41 cm

29

Presentation Player

RIJKS MUSEUM amsterdam

Powered by Cuypers

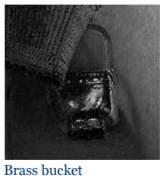
Duration: 3 minutes *Character:* Prosaic *Artefact:* The Kitchen Maid

Text: Audio: Speed:

Tiny points of light

Nail with shadow

Cracked Windowpane



Brass bucket

30

Flexible interfaces to MM

- Creating multimedia presentations requires
  - understanding message of presentation
  - knowing specifications of use context
  - making design dependencies explicit
  - taking these dependencies into account

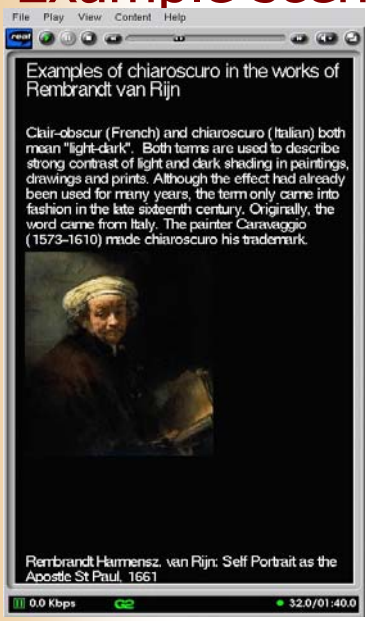
31

The Web in three generations

- 1 Hand-coded (HTML) Web content
  - easy access through uniform interface
  - huge authoring and maintenance effort
  - hard to deal with dynamically changing content
- 2 Automated on-the fly content generation
  - based on templates filled with database content
  - later extended with XML document transformations
- 3 Automated processing of content
  - The Semantic Web (SW)

32

Example scenario



Student is taking an art class on Rembrandt and wants to know about the "*chiaroscuro*" technique

System responds with a textual and audio explanation of the technique and a number of example images of its application in Rembrandt's paintings

33

2nd generation multimedia

- Adapt to end-user's platform capabilities
  - PC, PDA, mobile, voice-only, ...
- Adapt to the network resources available
  - bandwidth and other quality of service parameters
- Personalization
  - language, abilities, level of expertise, ..
- Problem: current 2nd generation Web tools **do not work for multimedia**

34

Multimedia differs from text

- Different document and presentation abstractions
  - hard to separate style from structure
- Formatting is not based on text flow
  - no pages or scrollbars, no line-breaking or hyphenation
  - templates often do not work well either
- Feedback from the formatting back-end required
  - need to check whether proposed layout is feasible
  - layout of media items is less flexible than text layout
- Transformations are hard in a functional language
  - need to try out designs and backtrack when necessary

35

Cuypers multimedia generation engine

- Cuypers web page http://media.cwi.nl/demo/
- Chiaroscuro example

- Acknowledgements:
  - Demonstrator developed in the context of the ToKeN2000 project
  - Media database used with permission, courtesy Rijksmuseum Amsterdam.

36

Cuypers – the bad news

The design knowledge is:

- implicit and hidden in the generation rules
- lost in the generated Web presentation
- not reusable for other Web applications/sites

We need the Semantic Web

37

Towards 3rd generation MM

38

Using an existing ontology

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="Painter">
      <rdfs:subClassOf rdf:resource="#Artist"/>
   </rdfs:Class>
   <rdfs:Class rdf:ID="Painting">
      <rdfs:subClassOf rdf:resource="#Artifact"/>
   </rdfs:Class>
...
</rdf:RDF>
```

39

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/…"
               xmlns:museum="http://ics.forth.gr/.../museum.rdf"

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

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

          <museum:Painting rdf:about="#apostlePaul">
            <museum:exhibited rdf:resource="#Rijksmuseum" />
            <museum:technique>chiaroscuro</museum:technique>
          </museum:Painting>
       </rdf:RDF>
    </metadata>
    . . .
```

40

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 MM DBMS..."/>
      <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">
          <img  region="img" src="..."/>
          <text region="ptitle" src=".."/>
        </par>
      </seq>
    </par>
  </body>
</smil>
```

41

Marked-up presentation


```
  <museum:Painting rdf:about="#apostlePaul">
        <museum:exhibited rdf:resource="#Rijksmuseum"/>
        <museum:technique>chiaroscuro</museum:technique
  >
  </museum:Painting>
   </rdf:RDF>
   </metadata>
...
  </head>
  <body>
    <par>
      <text region="title" src="..."/>
      <text region="descr" src="..."/>
      <seq>
        <par dur="10"> ... </par>
        <par dur="10"> ... </par>
        <par dur="10"> ... </par>
        <par dur="10"> ... </par>
        <par dur="10" id="apostlePaul">
          <img  region="img" src="..."/>
          <text region="ptitle" src=".."/>
        </par>
      </seq>
    </par>
  </body>
</smil>
```

This research is supported by

- NWO I²RP ToKeN2000
  Intelligent Information Retrieval and Presentation
- NWO CHIME ToKeN2000
  Cultural Heritage in an Interactive Multimedia Environment
- NWO NASH
  Networked Adaptive Structured Hypermedia
- Images courtesy of Rijksmuseum, Amsterdam

43

INRIA RHÔNE-ALPES

ina

Formalization of documentary knowledge and conceptual knowledge with ontologies :

applying to the description of audio-visual documents

Raphaël Troncy

Background

- The *audio-visual* document : some peculiarities
  - structured
  - spatio-temporal
  - composed of images

  } use of a textual description
- The *digital* audio-visual document :
  - allow new possibilities :
    - *« intelligent » search*
    - AV library structuring
    - re-use, publication and broadcasting
  - need for a hyper-linked description:
    the content has to be linked with the description

45

Plan of this talk

1. Problems
2. Document engineering *vs.*
   knowledge representation
3. Our proposal: an architecture for reasoning on descriptions of video documents
4. Conclusion and future work

46

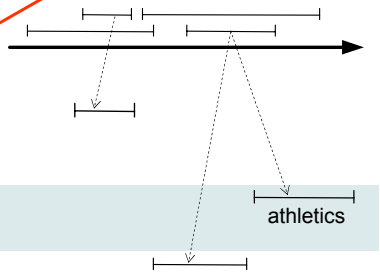
Description of the AV content

- A three step process :
  - ***identification*** of the content creator and the content provider : Dublin Core metadata, VRA core categories …
  - ***structural decomposition*** in video segments corresponding to the logical structure of the program : time-code, spatial coordinates
  - ***semantic description*** of these segments : controlled vocabulary, thesaurus, free text annotation

Description of the AV content

- Segmentation
  - locate and date some events
- Description
  - characterize each segment with an AV genre
  - characterize each segment with a general thematic
  - describe the scene (*who, when, where, what…*)

describe the logical structure



describe the semantics of the content

Example

13 [Indoor Set: 6th part]
at 18:43:56:00 - 00:09:06:00. – Eurosport
In studio, the second part of the interview, from Nice, of Sandy CASAR by Jean René GODART about the Paris-Nice cycling race and a few sports news with pictures commented by Alexandre BOYON and Laurent PUYAT.

Q : Find all AV sequences of type dialog sequence with a rider and concerning any *cycling race with several stages*

- noisy answer: there are other *sports news* in the sequence
- incomplete answer: the interview was broadcasted in two parts and began in a previous sequence
- the query cannot be extended !

Problems

- Weak use of the logical structures
- Descriptions are not made for reasoning

⇒ make the AV descriptions accessible to automated processes

⇒ What kind of knowledge do we need ?

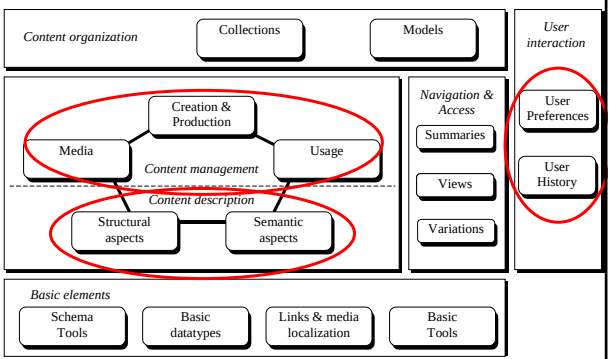
⇒ Which languages are the most suitable to perform all these tasks ?

Document engineering

- Provide models, languages and tools for managing document libraries
- Encode both structured documents and structured data: **XML** [W3C, 1998] & **XML Schema** [W3C, 2001]
- Distinguish the content from its presentation
  - Languages for presenting multimedia documents : **SMIL**
  - Models for describing multimedia documents
    - from **HyTime** [ISO, 1997] to **MPEG-7** [ISO, 2001]

MPEG-7, the new multimedia description language?

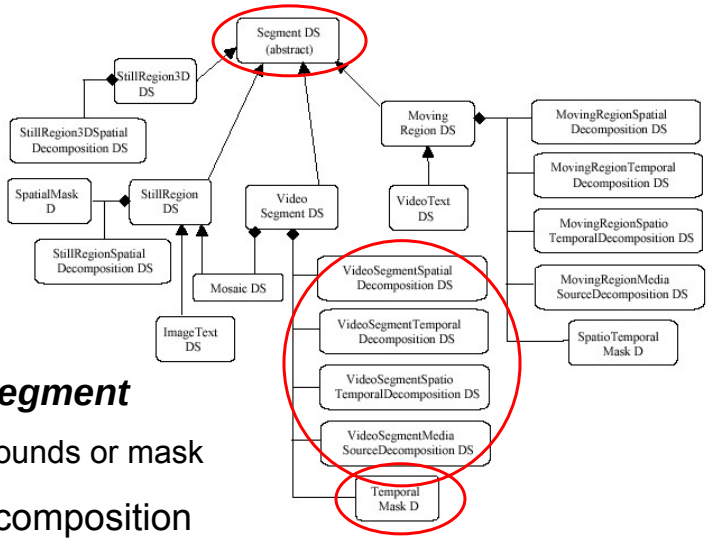
- ISO standard since December of 2001
- Main components:
  - Descriptors (Ds) and Description Schemes (DSs)
  - DDL (XML Schema + extensions)
- Concern all types of media



Part 5 - MDS

Structure and semantics

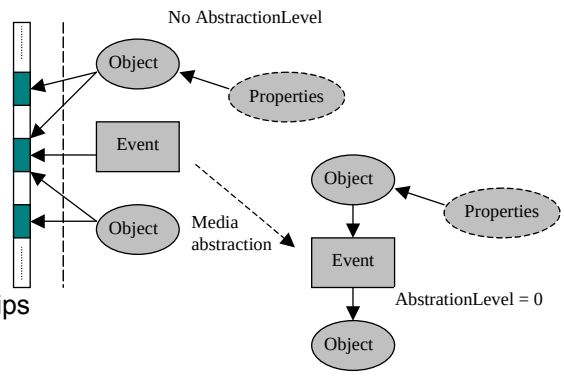
- Structure



- Base unit: ***segment***
  - temporal bounds or mask
- Possible decomposition

Structure and semantics

- Semantics
  - entity
  - attribute
  - relation
- *Classification Schemes* (CS)
  - thesaurus relationships



Other models

- MPEG-7 = a rich set of descriptors, but too restrictive to cover *all* the possible descriptions
- MPEG-7 extension with XML Schema:
  - Example: *TV Anytime*, *Mdéfi* [Tran Thuong, 2003]
  - Problem: add structure without semantics
- MPEG-7 extension with CS :
  - Example: the *COALA* system [Fatemi, 2003]
  - Problem: very poor expressivity
    (more specific, more general, related to, uses, is used by)

⇒ MPEG-7+XML Schema are not enough!
… but KR brings new solutions

Ontologies in KR

- The formal specification of a conceptual model for a given domain
  - A set of concepts, of relations and axioms
  - Knowledge representation languages
- Methodologies of construction:
  - Adaptation of well-known software engineering guidelines: *Methontology* [Gomez-Perez]
  - Terminological acquisition: [Bachimont], [Aussenac Gilles]
  - Ontology cleaning with formal properties: [Guarino]
- Tools :
  - *Protégé*, *WebODE*, *OilEd*, *OntoEdit*, *Terminae*, *DOE*

KR languages for the Web

- RDF : [W3C, 1999 & W3C, 2004]
  - a data model for annotating Web resources
  - triples: *resource* → *property* → *value*
- RDFS : [W3C, 2004]
-

```
<rdf:RDF>
   <ina:SportsNews rdf:about="Stade 2">
      <ina:broadChannel rdf:resource="France2"/>
      <ina:broadDate>17-03-2002</ina:broadDate>
   </ina:SportsNews>
</rdf:RDF>
```

```
(:"Stade 2" rdf:type ina:SportsNews)
(:"Stade 2" ina:broadChannel "France2")
(:"Stade 2" ina:broadDate 17-03-2002)
```

Use of OWL+RDF for describing AV documents

-
-
-

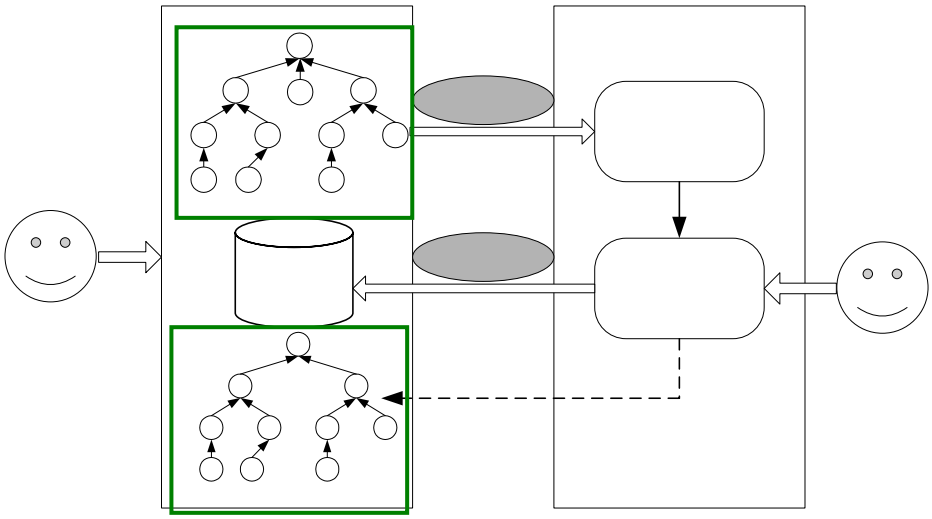
```
<owl:Class rdf:ID="TVProgram"/>

<owl:Class rdf:ID="StudioProgram">
  <rdfs:subClassOf rdf:resource="#TVProgram"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasPart"/>
      <owl:allValuesFrom rdf:resource="#StudioSequence"/>
    </owl:Restriction>
  </rdfs:subClassOf>
<owl:Class>

<owl:ObjectProperty rdf:ID="hasPart">
   <rdf:type rdf:resource="&owl;TransitiveProperty"/>
   <rdfs:domain rdf:resource="#TVProgram"/>
   <rdfs:range rdf:resource="#TVSequence"/>
</owl:ObjectProperty>
```

⇒ Problem: how to control the structure of the descriptions ?

Our proposition

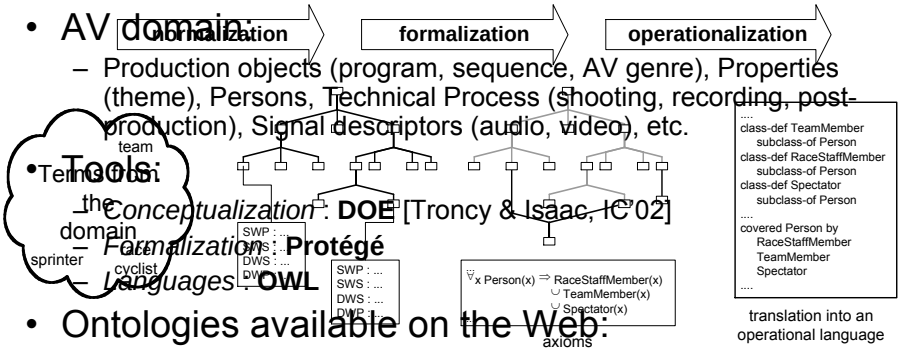
- Use jointly both approaches for representing the descriptions
  - the markup languages for describing and controlling the structure of each program
  - the ontology and the KR languages for describing formally the semantics of this structure and the content
- Automate the translation between these two representations as much as possible
- Develop a software architecture for reasoning on descriptions of video documents

General architecture

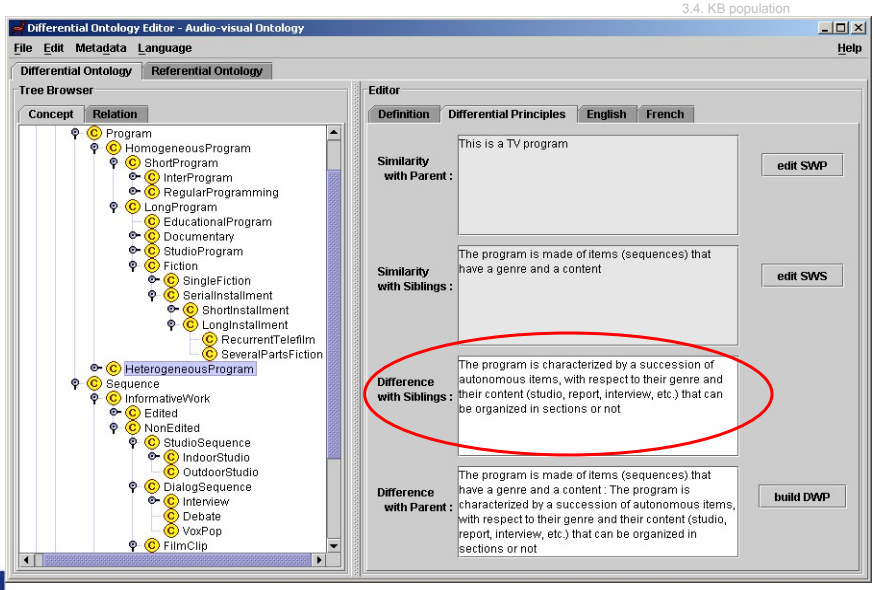


The Audio-visual Ontology

- Methodology of construction: **ARCHONTE** [Bachimont]
  - *Conceptualization* : differential principles
  - *Formalization* : formal definitions, axioms
  - *Operationalization* : export into a KR language
- AV domain:
  - Production objects (program, sequence, AV genre), Properties (theme), Persons, Technical Process (shooting, recording, post-production), Signal descriptors (audio, video), etc.
- Tools:
  - *Conceptualization* : **DOE** [Troncy & Isaac, IC'02]
  - *Formalization* : **Protégé**
  - *Languages* : **OWL**
- Ontologies available on the Web:
  http://www.cwi.nl/~troncy/ontologies/



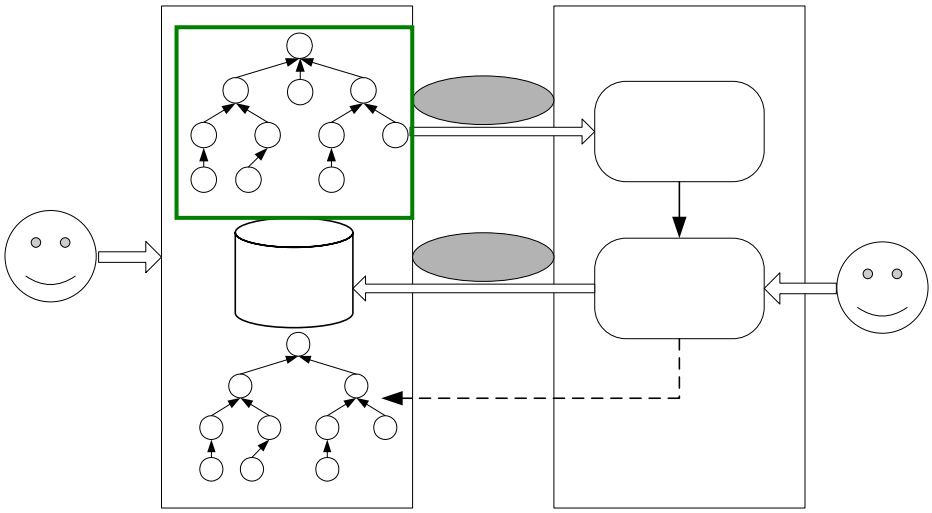
The *DOE* ontology editor



OWL Formalization


```
<owl:Class rdf:ID="TVProgram"/>

<owl:Class rdf:ID="StudioProgram">
  <rdfs:subClassOf rdf:resource="#TVProgram"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasPart"/>
      <owl:allValuesFrom rdf:resource="#StudioSequence"/>
    </owl:Restriction>
  </rdfs:subClassOf>
<owl:Class>

<owl:ObjectProperty rdf:ID="hasPart">
   <rdf:type rdf:resource="&owl;TransitiveProperty"/>
   <rdfs:domain rdf:resource="#TVProgram"/>
   <rdfs:range rdf:resource="#TVSequence"/>
</owl:ObjectProperty>
```


General architecture



Generate *XML Schema* types

Some concepts (*program*, *sequence*) refer to categories of audio-visual segments

OWL

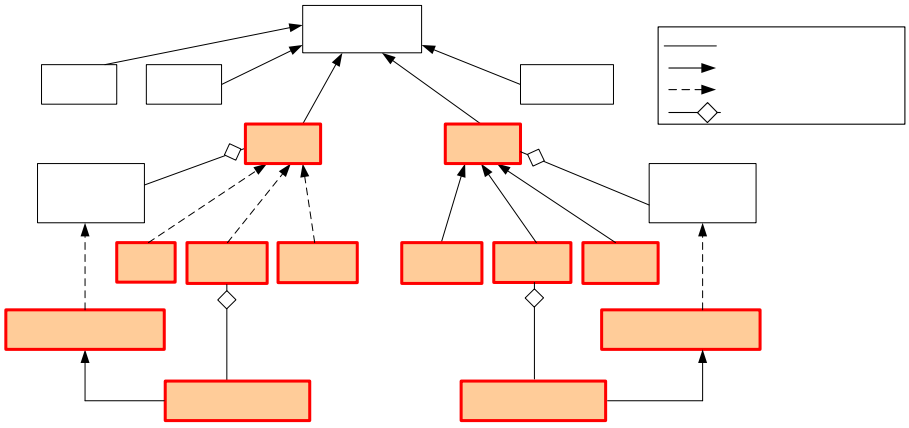
- Class
- Sub-class
- Restriction on properties
- Union of classes

XML Schema

- Complex type
- Extension
- Element of the content model
- Choice in the content model

Generic MPEG-7 extension

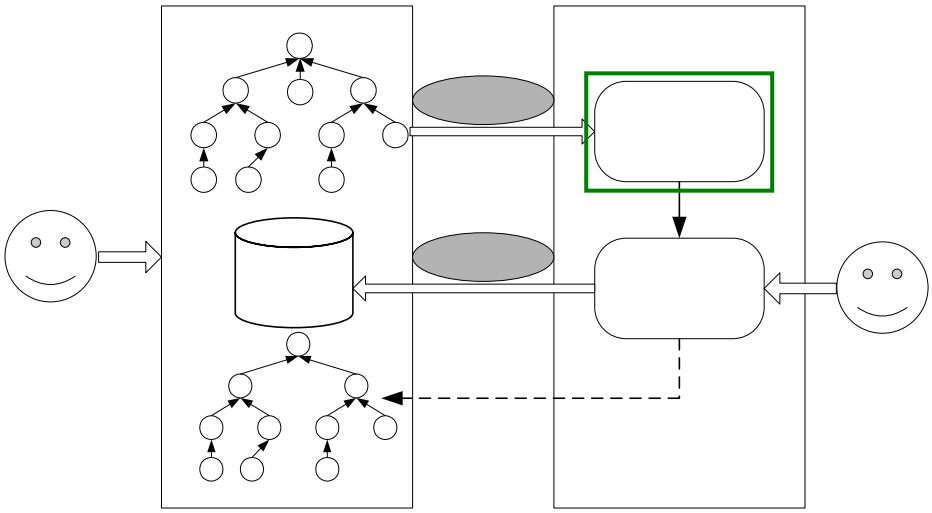
- Link these types to the existing MPEG-7 types



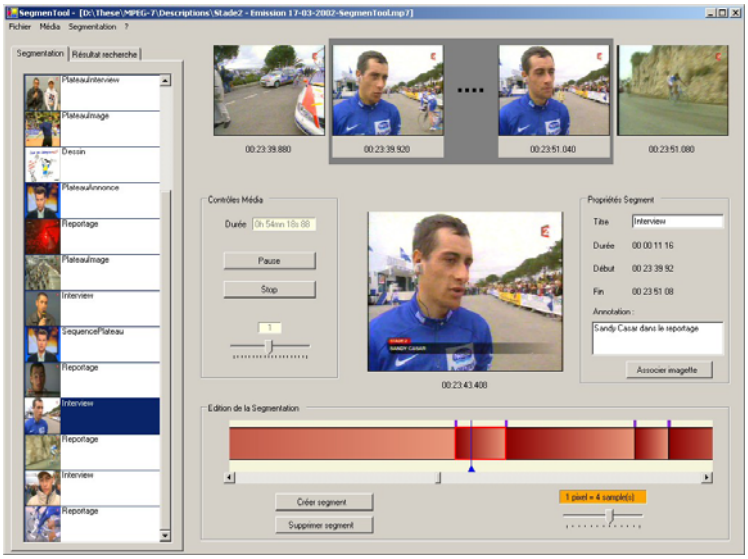
Build description schemes

- Let us watch some sports magazines
  - construction of a simple schema based on *StudioSequence*, *Report* and *Interview*
  - a *Report* contains some *Excerpts* of *Broadcast Live Sports*
- The schema provides the description skeleton for several sports magazine:
  - **Téléfoot** (soccer)
  - **VéloClub** (cycling)
  - **3 Partout** (multisports)

General architecture

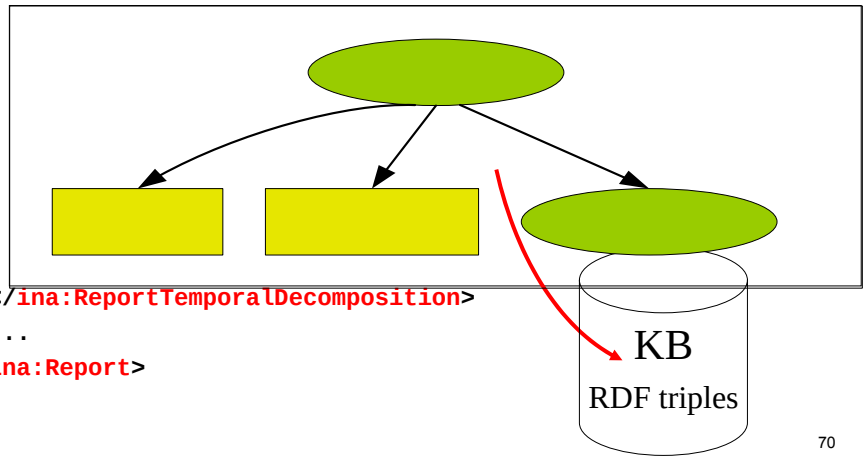


SegmenTool [French projet CHAPERON]



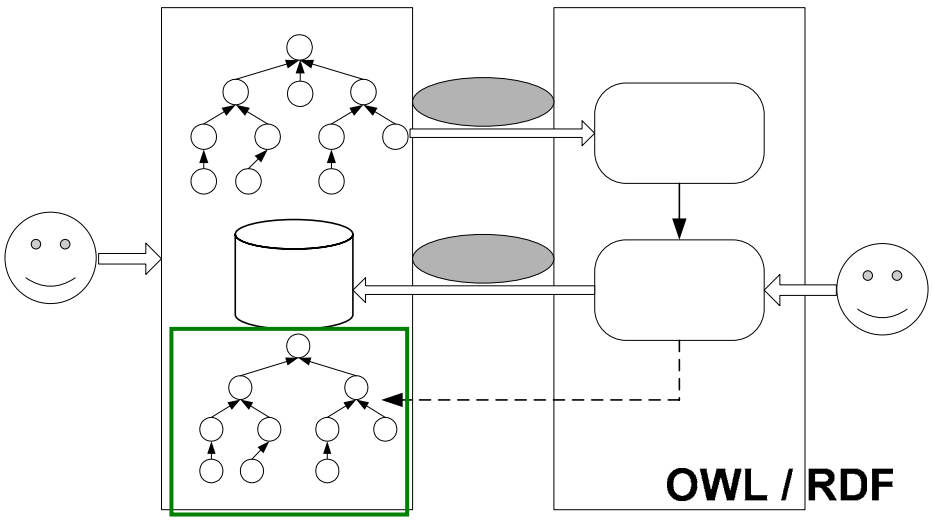
Instantiate a document content model


```
<ina:Report id="aa23c647c-6517-4aee-8bce-870ae52a01af">
  ...
  <ina:ReportTemporalDecomposition>
```




```
  </ina:ReportTemporalDecomposition>
  ...
</ina:Report>
```

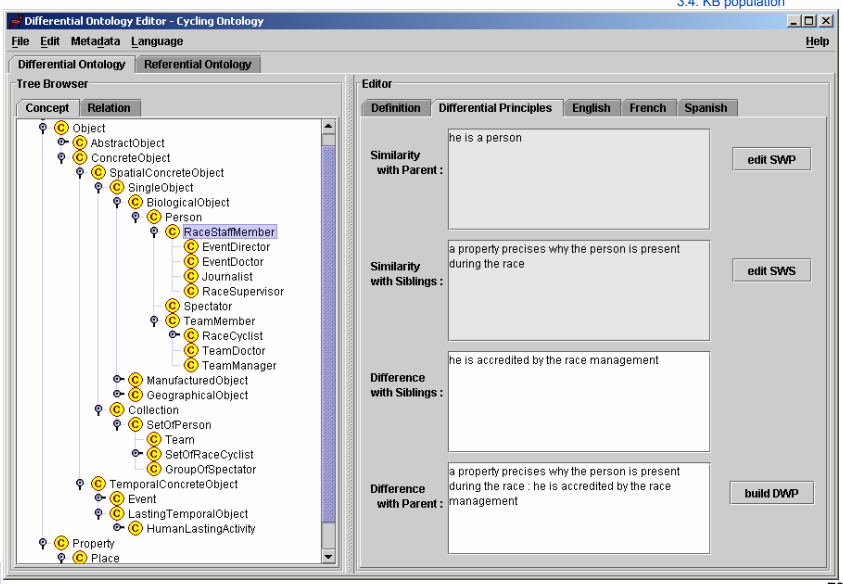
General architecture



The Cycling Ontology

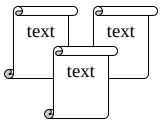
- Methodology of construction:
  - Terminological acquisition
    - Textual corpus of 550 000 words [LeRoux, 2003]
    - Tool for candidate term extraction: *Lexter*
  - Conceptualization and formalization
    - *DOE + OilEd*
- Results:
  - Construction time: 3 weeks
    - conceptualization, upper level, formalization
  - Ontology size: medium
    - 97 concepts, 61 relations

The Cycling Ontology



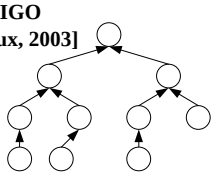
Knowledge Base population

Cycling domain



Base of facts

SEIGO
[Le Roux, 2003]



```
<rdf about="{URI}/MagazineSportif5/Report3/Interview4">
   <!- formal statements from a base of facts -->
</rdf>
```

Conclusion

- General architecture for reasoning on descriptions of video documents:
  - Control of the structure: creation of document schemes
  - Formal representation of the semantics: AV ontology and domain-specific ontology
  - Based on standards languages (MPEG-7, OWL, RDF) and the use of transformations
- Implementation
  - Generic extension of MPEG-7
  - Modeling of 2 ontologies with ***DOE***
  - Creation of a Knowledge Base of events related to cycling race and use of an adapted reasoner

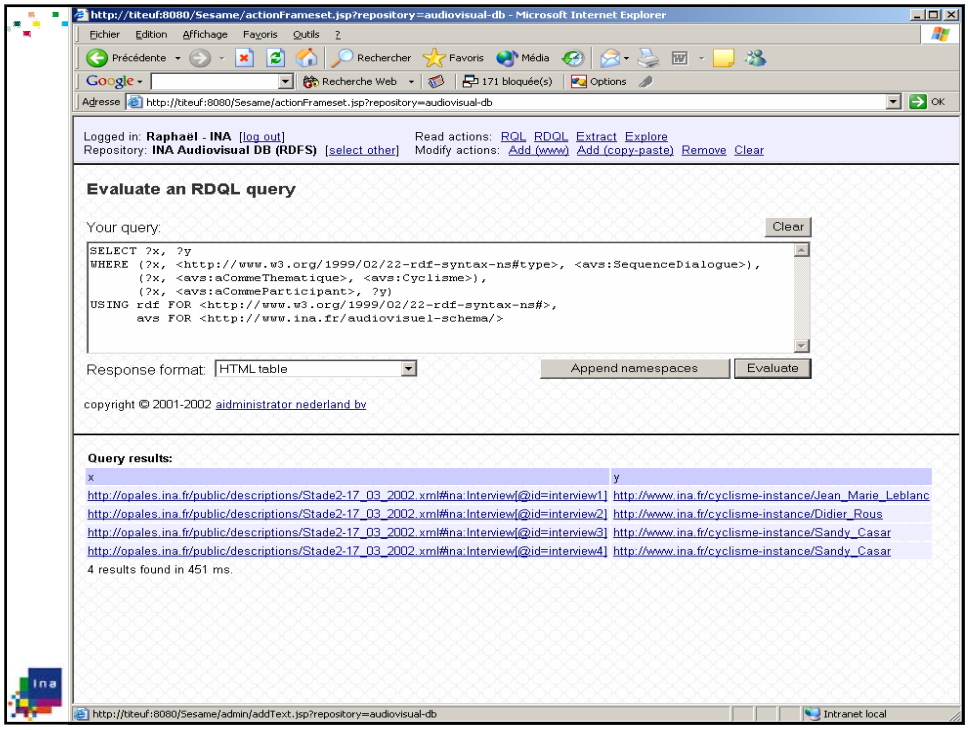
Future work

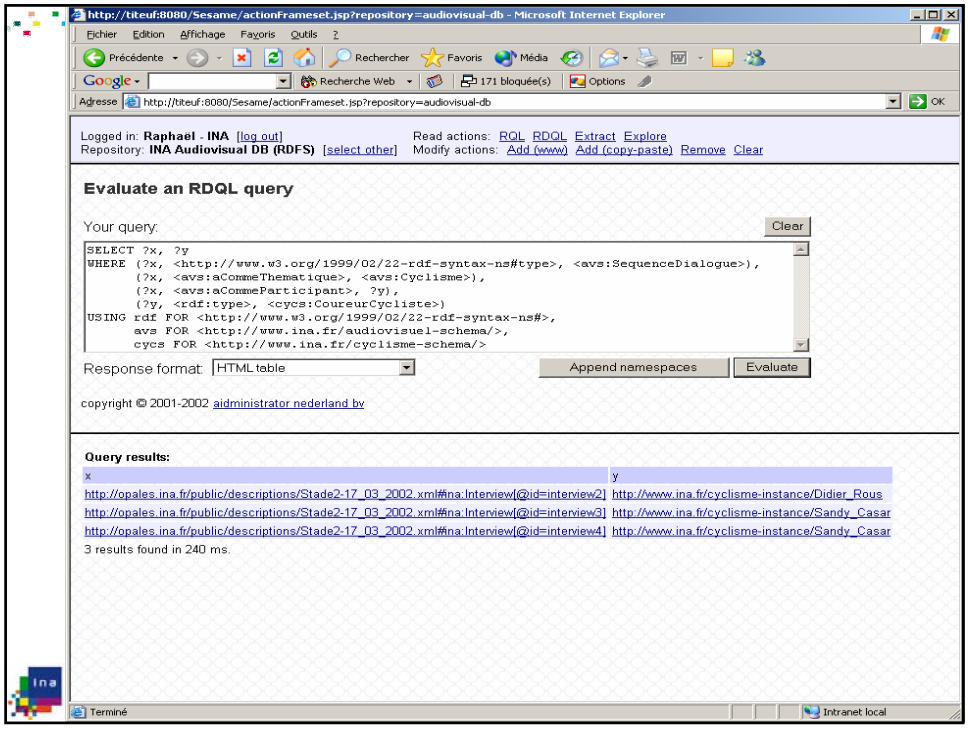
- Long-term objectives
  - The *ideal* AV description language is still a research program
  - The description could be linked with:
    - a *rhetorical* analysis of the documents
    - a *semiotic* analysis of the documents

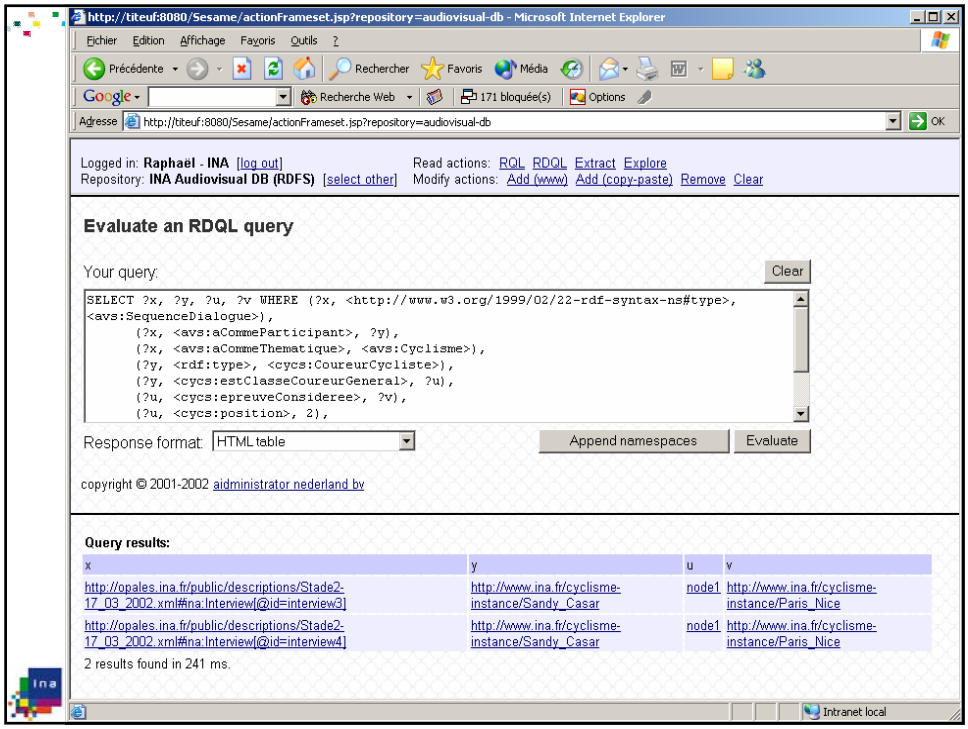
Questions?

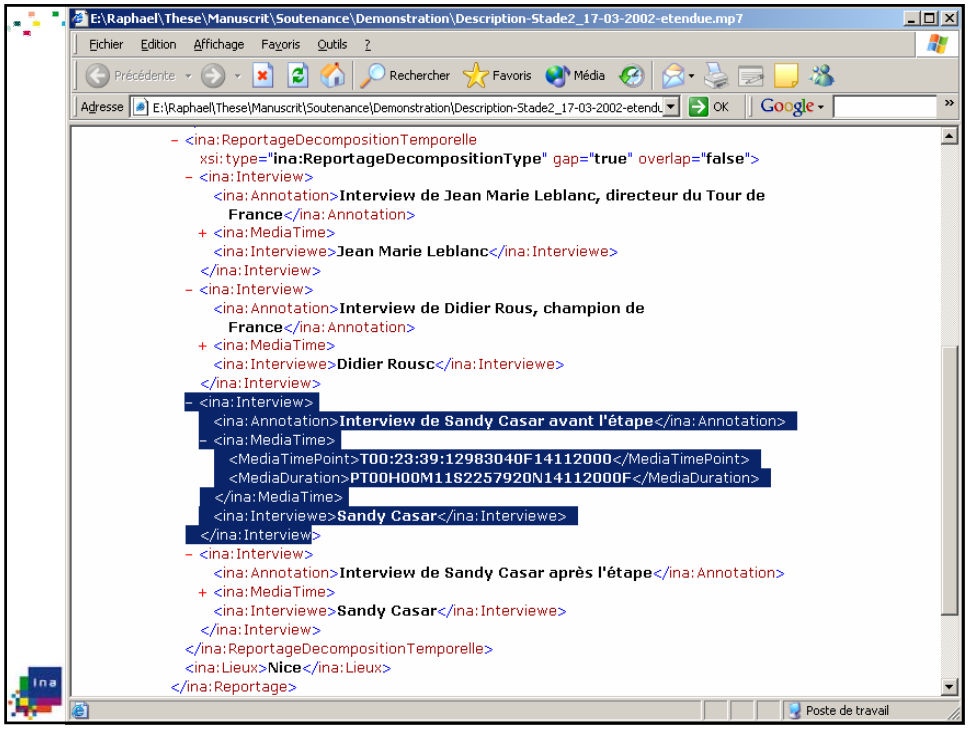
1. Problems
2. Document engineering *vs*. knowledge representation
3. Our proposal: an architecture for reasoning on descriptions of video documents
4. Conclusion and future work

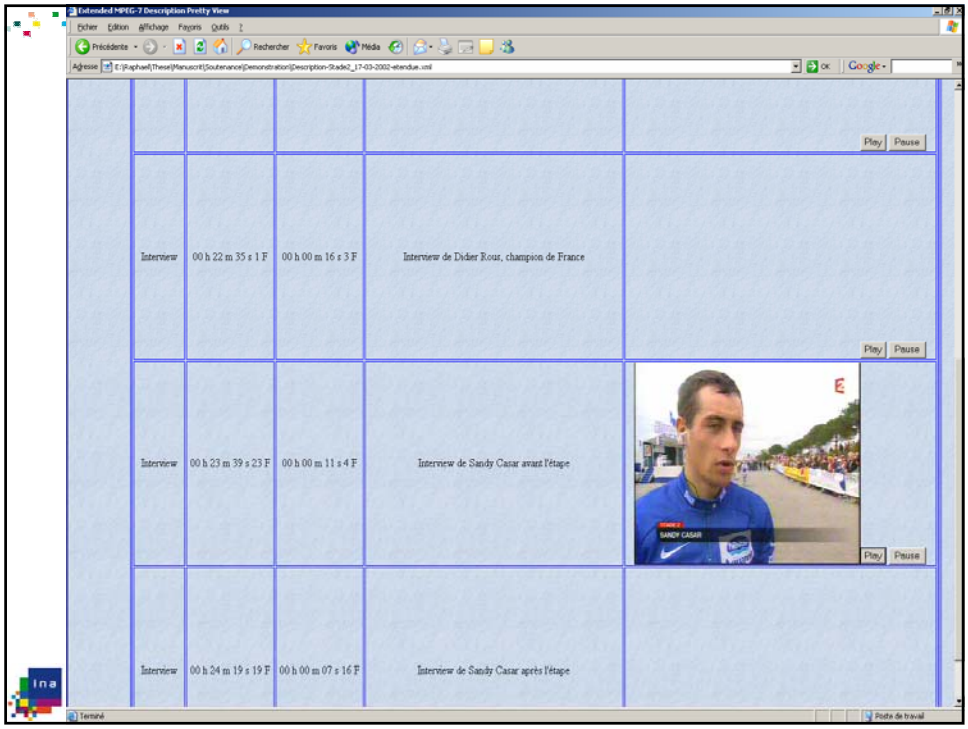
77

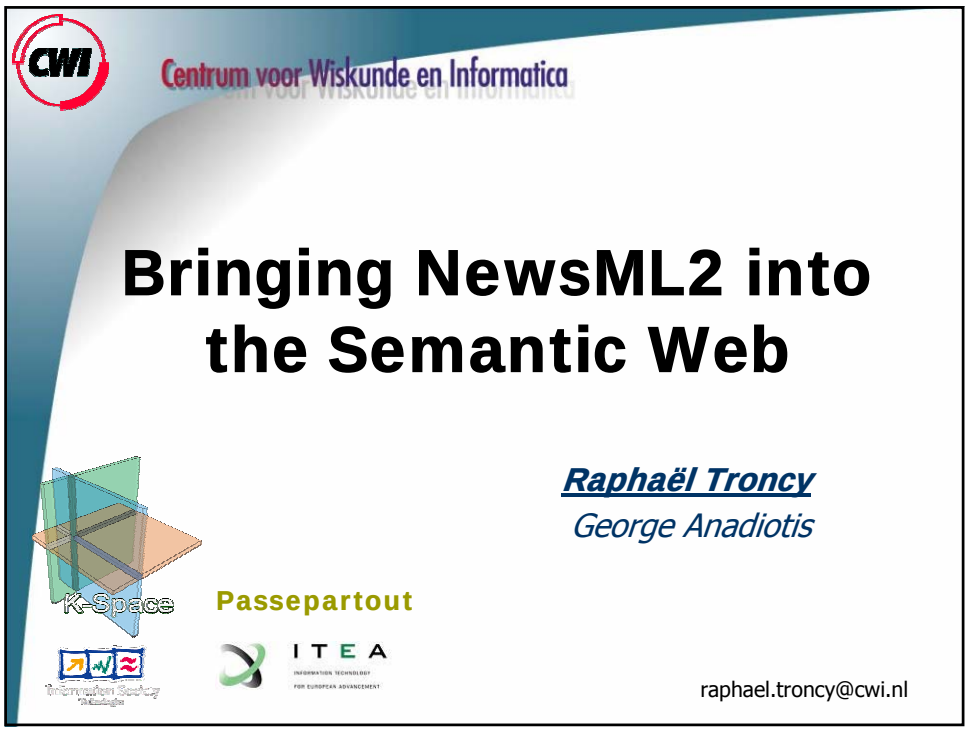



http://titeuf:8080/Sesame/actionFrameset.jsp?repository=audiovisual-db - Microsoft Internet Explorer
Logged in: Raphaël - INA [log out]
Repository: INA Audiovisual DB (RDFS) [select other]
Read actions: RQL RDQL Extract Explore
Modify actions: Add (www) Add (copy-paste) Remove Clear
Evaluate an RDQL query
Your query:
Clear
SELECT ?x, ?y
WHERE (?x, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>, <avs:SequenceDialogue>),
(?x, <avs:aCommeThematique>, <avs:Cyclisme>),
(?x, <avs:aCommeParticipant>, ?y),
(?y, <rdf:type>, <cycs:CoureurCycliste>)
USING rdf FOR <http://www.w3.org/1999/02/22-rdf-syntax-ns#>,
avs FOR <http://www.ina.fr/audiovisuel-schema/>,
cycs FOR <http://www.ina.fr/cyclisme-schema/>
Response format: HTML table
Append namespaces
Evaluate
copyright © 2001-2002 aidministrator nederland bv
Query results:
x
y
http://opales.ina.fr/public/descriptions/Stade2-17_03_2002.xml#na:Interview[@id=interview2]
http://www.ina.fr/cyclisme-instance/Didier_Rous
http://opales.ina.fr/public/descriptions/Stade2-17_03_2002.xml#na:Interview[@id=interview3]
http://www.ina.fr/cyclisme-instance/Sandy_Casar
http://opales.ina.fr/public/descriptions/Stade2-17_03_2002.xml#na:Interview[@id=interview4]
http://www.ina.fr/cyclisme-instance/Sandy_Casar
3 results found in 240 ms.


http://titeuf:8080/Sesame/actionFrameset.jsp?repository=audiovisual-db - Microsoft Internet Explorer
Logged in: Raphaël - INA [log out]
Repository: INA Audiovisual DB (RDFS) [select other]
Read actions: RQL RDQL Extract Explore
Modify actions: Add (www) Add (copy-paste) Remove Clear
Evaluate an RDQL query
Your query:
Clear
SELECT ?x, ?y, ?u, ?v WHERE (?x, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>,
<avs:SequenceDialogue>),
(?x, <avs:aCommeParticipant>, ?y),
(?x, <avs:aCommeThematique>, <avs:Cyclisme>),
(?y, <rdf:type>, <cycs:CoureurCycliste>),
(?y, <cycs:estClasseCoureurGeneral>, ?u),
(?u, <cycs:epreuveConsideree>, ?v),
(?u, <cycs:position>, 2),
Response format: HTML table
Append namespaces
Evaluate
copyright © 2001-2002 aidministrator nederland bv
Query results:
x
y
u
v
http://opales.ina.fr/public/descriptions/Stade2-17_03_2002.xml#na:Interview[@id=interview3]
http://www.ina.fr/cyclisme-instance/Sandy_Casar
node1
http://www.ina.fr/cyclisme-instance/Paris_Nice
http://opales.ina.fr/public/descriptions/Stade2-17_03_2002.xml#na:Interview[@id=interview4]
http://www.ina.fr/cyclisme-instance/Sandy_Casar
node1
http://www.ina.fr/cyclisme-instance/Paris_Nice
2 results found in 241 ms.


E:\Raphael\These\Manuscrit\Soutenance\Demonstration\Description-Stade2_17-03-2002-etendue.mp7
Fichier Edition Affichage Favoris Outils ?
Précédente Rechercher Favoris Média
Adresse E:\Raphael\These\Manuscrit\Soutenance\Demonstration\Description-Stade2_17-03-2002-etendu
OK Google
<ina:ReportageDecompositionTemporelle xsi:type="ina:ReportageDecompositionType" gap="true" overlap="false">
<ina:Interview>
<ina:Annotation>Interview de Jean Marie Leblanc, directeur du Tour de France</ina:Annotation>
<ina:MediaTime>
<ina:Interviewe>Jean Marie Leblanc</ina:Interviewe>
</ina:Interview>
<ina:Interview>
<ina:Annotation>Interview de Didier Rous, champion de France</ina:Annotation>
<ina:MediaTime>
<ina:Interviewe>Didier Rousc</ina:Interviewe>
</ina:Interview>
<ina:Interview>
<ina:Annotation>Interview de Sandy Casar avant l'étape</ina:Annotation>
<ina:MediaTime>
<MediaTimePoint>T00:23:39:12983040F14112000</MediaTimePoint>
<MediaDuration>PT00H00M11S2257920N14112000F</MediaDuration>
</ina:MediaTime>
<ina:Interviewe>Sandy Casar</ina:Interviewe>
</ina:Interview>
<ina:Interview>
<ina:Annotation>Interview de Sandy Casar après l'étape</ina:Annotation>
<ina:MediaTime>
<ina:Interviewe>Sandy Casar</ina:Interviewe>
</ina:Interview>
</ina:ReportageDecompositionTemporelle>
<ina:Lieux>Nice</ina:Lieux>
</ina:Reportage>
Poste de travail


Play Pause
Interview 00 h 22 m 35 s 1 F 00 h 00 m 16 s 3 F Interview de Didier Rous, champion de France
Play Pause
Interview 00 h 23 m 39 s 23 F 00 h 00 m 11 s 4 F Interview de Sandy Casar avant l'étape
Play Pause
Interview 00 h 24 m 19 s 19 F 00 h 00 m 07 s 16 F Interview de Sandy Casar après l'étape



Why Bother with Metadata?

- A News agency is a content provider
  - Content (stories, photo, video, etc.) are assets
- Metadata add value to these assets as they provide human and machine readable information about them
- Metadata is much more than just a bunch of keywords added at the end of the chain so the customer can find your image
- Metadata covers all information about an asset, which enables machines to do smart things with your assets

84

Why Bother with Semantics?

- High quality *semantic* multimedia metadata enables:
  - Easy exchange of news items
  - Semantic search of particular news items
  - Delivery of personalized news content to customers
    - Interactive browsing in a news archive
    - Cross-modality: packaging the news stories, photos, graphics, audio, videos
    - For different end-user platforms (mobiles, PC, handhelds, etc.)

85

IPTC Metadata Standards

- Metadata "fields"
  - Informal definition and guidelines to use the field according to its semantics
  - e.g. "Date Created": content creation date ≠ digital representation creation date

Property name: **Creator**
User interface label: Creator
Description: Contains preferably the name of the person who created the content of this news object, a photographer for photos, a graphic artist for graphics, or a writer for textual news. If it is not appropriate to add the name of a person the name of a company or organisation could be applied as well.
Note(s): Aligning with IIM notions IPTC Core intents to have only one creator for this news object despite the underlying XMP property dc:creator allows for more than one item to be included. If there are more than one item in this array the first one should be considered as the IPTC Core Creator value.
XMP Schema specifications:
XMP Category: External XMP Value Type: Seq ProperName
XMP Path: dc:creator/*[1]

86

IPTC Metadata Standards

- Metadata "values"
  - Expressed as *controlled* vocabularies (standardization bodies)
  - A vocabulary is composed of terms (flat list, taxonomy organization)
  - IPTC has defined 28 sets of multilingual News Codes
    - NewsCodes use numeric strings = language agnostic
    - Ex: Subject ≈ 1300 terms, 3 levels hierarchy in 4 languages
    - NewsCodes Viewer application View
- XML Wrapper
  - Metadata embedded in a photo: XMP
  - Metadata stored in a separate file: NewsML

Problem: XML and Semantic *)

うかを検出するために、文書の完全性を保証することです。しかしながら多くのアプリケーションは、XML 文書にまず署名をし、その後文書を改変することで、その文書の一部を暗号化しようと考えています。復号化変換では、署名の確認に先立ち、文書を改変前の状態に戻し、文書のどの部分を復号化すればよいかをデータ受信者に通知します。

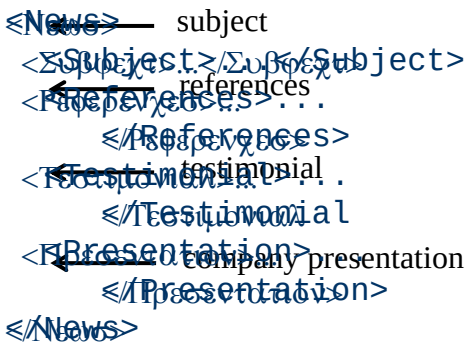
業界リーダーや暗号の専門家らの幅広い支持とともに、既に実装もされている XML Encryption

W3C の XML Encryption ワーキンググループによってまとめられた実装及び相互運用性報告書に示されているように、数多くのアプリケーションや他の仕様が既に XML Encryption を利用しています。特に、配送データのセキュア化が必要な Web サービス仕様群が本仕様の利用を進めています。また多くの企業が XML Encryption の実装についてその支持と計画を表明しています。

XML Encryption は、Baltimore Technologies、BEA Systems、DataPower、IBM、Microsoft、Motorola、ジーゲン大学、Sun Microsystems、VeriSign の各 W3C 会員と個人技術者とで構成される、W3C の XML Encryption ワーキンググループによって策定されました。

World Wide Web Consortium [W3C] について

W3C は、Web の発展と相互運用性を確保するための共通のプロトコルを開発することにより、Web の可能性を最大限に引き出すべく設立されました。W3C は、アメリカ合衆国マサチューセッツ工科大学計算機科学研究所 (MIT/LCS)、フランス国立情報処理自動化研究所 (INRIA)、及び日本の慶應義塾大学がホスト機関として共同運営にあたっている国際産業コンソーシアムです。コンソーシアムにより提供されるサービスには、開発者及び利用者のための World Wide Web に関する豊富な情報、新技術を応用した様々なプロトタイプやサンプルアプリケーションの開発などが挙げられます。現在までに、450近くの組織がコンソーシアムの会員となっています。詳しくは http://www.w3.org/ をご参照下さい。



⇒ **Need for formal semantics for the content**

*) adapted from Frank van Harmelen

Problem: interoperability

- Different management applications may label the same field differently
  - e.g. Creator / By-Line (Author) / Author / By-Line
- The informal semantics (guidelines) of the various metadata fields prevent an automatic validation of their use

⇒ **Need for formal semantics for the structure**

89

Role of the Semantic Web

- "Oh no! Not yet another metadata standard!" Like we don't have enough of them already:
  - EXIF, Dublin Core, VRA Core, IPTC Core, XMP, MPEG-7, Creative Commons, ... ?
- But again: No single standard can cover all metadata needs
- SW is a framework that could make existing metadata standards and tools interoperable ... and make them interoperable with the rest of the Web!

90

NewsML2 and the SW

- Common basis
  - Distributed resources (news item) globally and uniquely identified => URI
  - Use of shared and controlled vocabularies
- Natural switch and numerous benefits
  - Better control of NewsML2 descriptions (logical consistency check)
  - Enhanced search of News topic (logical inferences)
  - Intelligent presentation – Semantic interfaces
  - Unified news management – Semantic CMS

91

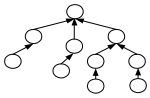
Use Case scenario

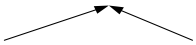
```
<newsItem schema="0.7" version="2">
  ...
  <itemMeta>
    <contentClass code="ccls:photo" />
    ...
  </itemMeta>
  <contentMeta>
    <infoSource literal="AFP" />
    <locCreated code="city:Kathmandu">
      <broader code="ctry:NEP" />
    </locCreated>
    <subject code="cat:01001000" type="ctyp:politics">
      <title>King</title>
    </subject>
    <description>
Nepal's King Gyanendra attended a Hindu festival in Kathmandu, his first
public appearance since being stripped of most of his powers by
parliament last month.
    </description>
  </contentMeta>
  ...
</newsItem>
```

Use Case scenario

Q: News about the *leader* of the *country Nepal* ?

The King Gyanendra of Nepal





Head State ⇔ and (King
(oneOf country Nepal, NL, ...))
Head Government ⇔ and (Prime Minister
(oneOf country Nepal, NL, ...))

The Prime Minister
Girija Prasad Koirala

93

What we have done?

- Creation of a News domain ontology in OWL
  - Based on the UML model specifications of NewsML2
- Online conversion service
  - Mapping of the IPTC NewsCodes into various SKOS thesaurus
  - Transforming dynamically the NewsML2 (XML) descriptions in its equivalent RDF counterpart
    - Using to the NewsML ontology
    - Linking to the SKOS IPTC NewsCodes

http://newsml.cwi.nl/

94

What is the added value?

- Example: A "normal" day in AFP
- Dataset
  - 200 NewsML2 stories, 35 photos (original size + thumbnails) + 35 NewsML2 descriptions
  - Covering various subjects:
    - A military drill for dealing with contaminations (toxic, nuclear or biological) - Photo
    - A regular meeting of the French cabinet - Photo
    - A strike in New Caledonia - Photo
    - A protest made on the Arch of Triumph in Paris, related to the Iran nuclear crisis - Photo
    - A wine makers protest - Photo
    - A meeting between the French president and Israeli prime minister - Photo
    - A senator's publicity pictures - Photo

95

Example 1: reasoning on the content

- Find all related news about "Nuclear"

Nucléaire → Military drill (NBC)

Nuclear → Iran nuclear crisis → *Arc de Triomphe* protest

Iran nuclear crisis → Chirac – Elmer summit

96

Example 2: reasoning on the structure

- Find photos of Y for which the author is X ?
- What the NewsML ontology provide ?
  - *slugline* and *headline* are *metadata properties*, whose values are *Basic Components*
  - *creator* and *contributor* are *authors*
  - history of the description (versioning)
- No need to know the NewsML structure to answer the query

97

What to do with the RDF data?

- Various tools that are able to digest RDF data and provide a unified view of these data
  - FOAF Viewer http://xml.mfd-consult.dk/foaf/explorer/
  - SIMILE project http://simile.mit.edu/piggy-bank/
- /facet: A Browser for Heterogeneous Semantic Web repositories
  - Faceted browser paradigm (*Flamenco*)
  - Provide a view on any RDF dataset

98

Conclusion

- Methods and conversion tools for bringing NewsML in the SW (RDF - compliant)
- Added-value:
  - Enhance search of news items (logical inferences on the structure and the content)
  - Enhance presentation of news items
    - Semantic media interfaces
    - Discover relations between Items / Topics / Packages
  - Semantic Content Management System
    - Keep track of provenance information

99

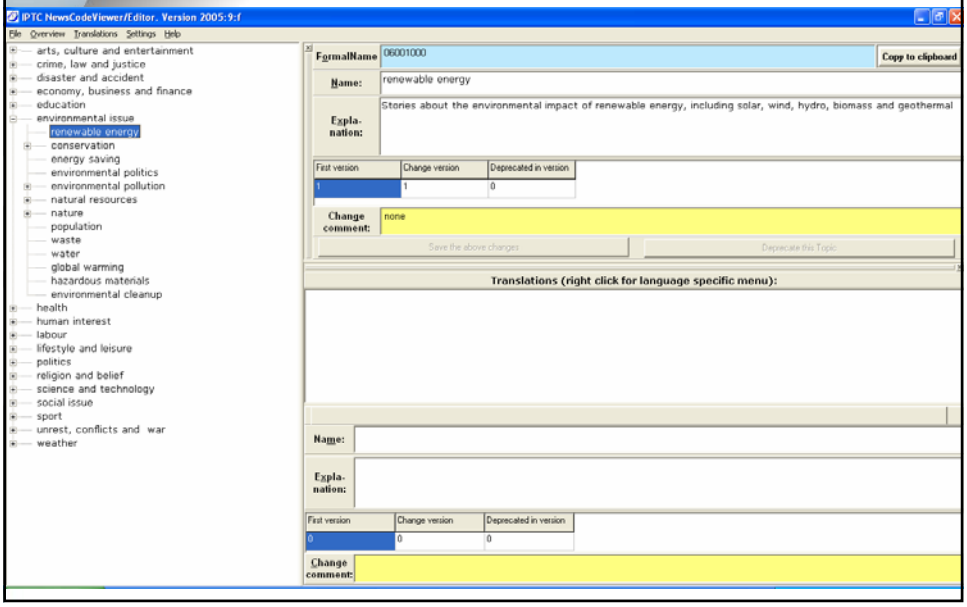
Future Work

- Making the use case scenario REAL!
  - Needs data: photos, videos, graphics, audio, textual stories !
- Implement interfaces for:
  - Browsing a News archive
  - Rendering the search results
- Establishing links between NewsML and other vocabularies
  - IPTC News Codes *versus* domain ontologies
  - NewsML *versus* DC, EXIF, MPEG-7, etc.

100

NewsCodeViewer

back



Myths about the Semantic Web *)

1. "SW people try to enforce meaning from the top"
   - They only recommend languages that you can use to define *your* concepts according to *your* definitions
2. "SW people will require everybody to subscribe to a single predefined 'meaning' for the terms we use"
   - You can use these languages to relate existing concepts (bridging communities)
3. The SW will require users to understand the complicated details of formalized knowledge representation
   - All of this 'under the hood'
4. "SW people will require us to manually annotate all the existing web-pages"
   - SW languages can be used to exchange manually *and* automatically produced metadata

102

*) adapted from Frank van Harmelen, WWW2006 panel "Meaning on the Web: Evolution or Intelligent Design?"

Generating video documentaries from annotated media repositories

Stefano Bocconi
CWI Amsterdam
The Netherlands
Contact: Stefano.Bocconi@cwi.nl

Talk Outline

- Motivation
- Example
- Scenarios
- Technical details
  - Annotations
  - Editing Process
- Conclusions

104

Video Documentaries on the Web

- Traditional video authoring: there is only one final version, what is shown is the choice of the author/editor
- Proposed video authoring:
  - Annotate the video material semantics
  - Show automatically what the user asks to see, using presentation forms a film editor would use

105

Video material

- Focus on video interviews about controversial issues
- **Interview with America** video footage with interviews and background material about the opinion of American people after 9-11
  www.interviewwithamerica.com

106

Example: *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…"

What do you think of the war in Afghanistan?

I am not a fan of military actions

I cannot think of a more effective solution

War has never solved anything

Two billions dollar bombs on tents

Scenarios

- Augmenting one interview with man-on-the-street opinion (“Vox Populi” documentary)
- Overview of the content of video footage:
  - Example: trailers (“Voices of Iraq”)
  - Browse the content by opinion

109

The annotations

- Rhetorical
  - Rhetorical Statement (mostly verbal, but visual also possible)
  - Argumentation model: Toulmin model
- Descriptive
  - Question asked
  - Interviewee (social)
  - Filmic (e.g. location/time/framing/gaze)

110

Encode statements

- Statement formally annotated:
  - <subject> <modifier> <predicate>
  - E.g. "**war best solution**"
- A thesaurus containing:
  - Terms (155)
  - Relations between terms: *similar* (72), *opposite* (108), *generalization* (10), *specialization* (10)
  - E.g. **war** *opposite* **diplomacy**

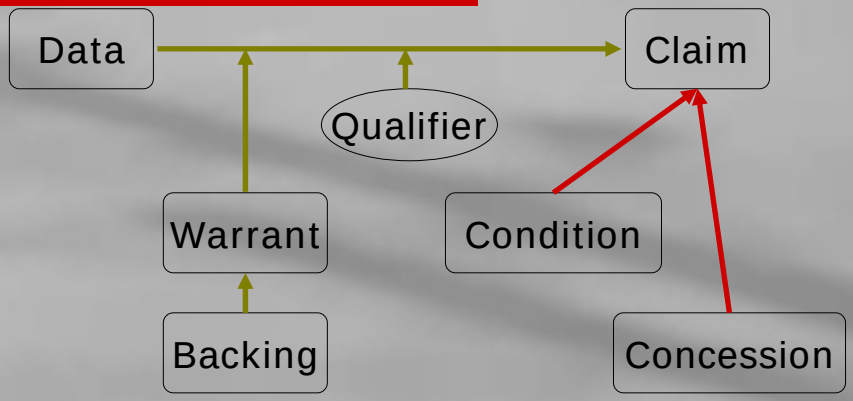
111

Connect statements

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

112

Toulmin model



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

113

Analysis of the Example

Two billions dollar bombs on tents

Claim

contradict

Claim

I cannot think of a more effective solution

weaken

Concession

I am not a fan of military actions

Claim

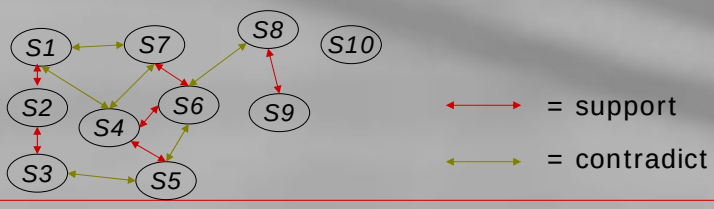
support

War has never solved anything

114

Facts and features

- Annotations: 1 hour annotated, 15 interviews, 60 interview segments, 120 statements
- Partially tunable: examining the Segment graph gives feedback on the quality of the annotations and the thesaurus



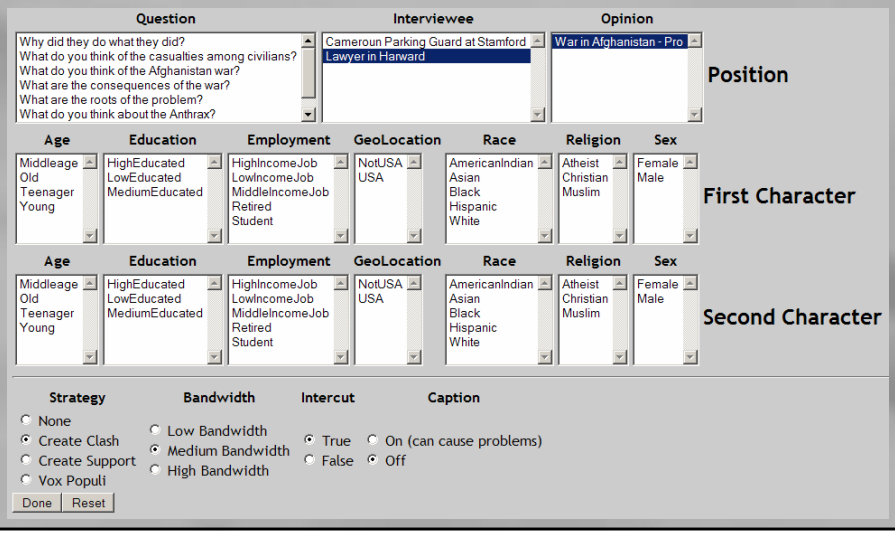
115

Controlling the Bias

- Video documentaries are not neutral account of reality: the selection and editing of the footage expresses a point of view
- Editing strategy:
  - Balanced
  - Pro opinion X
  - Against opinion X
- We use:
  - Logos (the statements)
  - Ethos (based on user profile)
  - Film editing (e.g. framing, gaze)

116

Vox Populi interface



Conclusions

- Automatic generation of video interviews augmented with supporting and/or contradicting material
- The **user** can determine the subject and the bias of the presentation
- The **documentarist** can add material and let the system generate new documentaries

Pointers & Acknowledgments

- This presentation and a Demo available at:
  http://www.cwi.nl/~media/demo/IWA/
- This research was funded by the Dutch national ToKeN I²RP and CHIME projects.

119

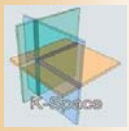
TU/e technische universiteit eindhoven

CWI

Creating Meaningful Presentations

Lynda Hardman
Jacco van Ossenbruggen

Semantic Media Interfaces
CWI, Amsterdam
http://www.cwi.nl/~media



Introduction

- Overview of our research activities:
  - Creating meaningful presentations from query results
  - Part of the K-Space, Passepartout and Multimedia-N
- Main theme of our work:
  - The role that ***explicit discourse*** information plays in the *personalized generation process*
  - The difference between:
    - a *list* of retrieval results ordered *most relevant first* and
    - a *presentation* that has *structure* interpretable by the end user, giving the collection *sense of belonging to same presentation*

121

Existing approaches in presenting query results

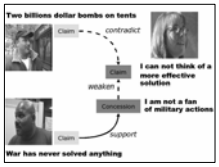
- No explicit discourse (only domain semantics):
  - Noadster - clusters from domain semanitcs
  - Topia - preselecting concepts in domain semantics
  - Museo Suomi - selection based on domain semantics
- Deriving some aspects of discourse:
  - Giving meaning to clustering process
  - Assigning different weights to clusters => ordering
    - Influence the way people perceive information

122

Explicit Discourse

- **Fixed discourse**:
  - DICS – uses annotated multimedia repository + domain ontology and discourse knowledge
  - discourse knowledge = set of rules (genre, narrative units…)

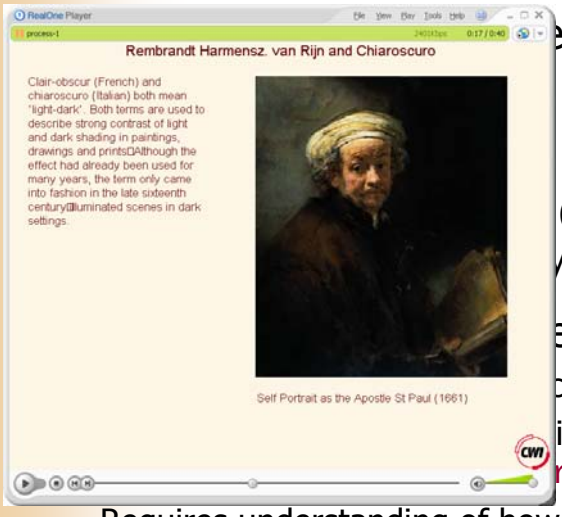
- **Dynamic discourse**:
  - VoxPopuli: argument generation in video

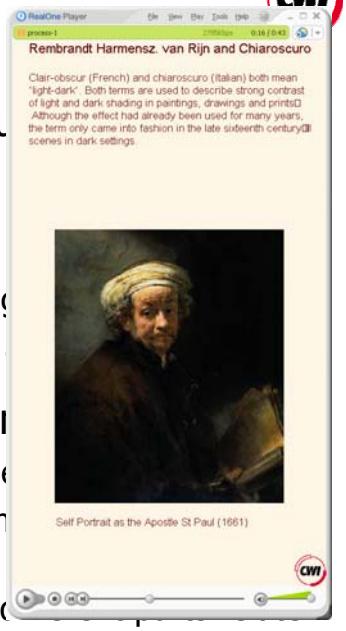


- **Role of structured progression**

123

Fixed Discourse



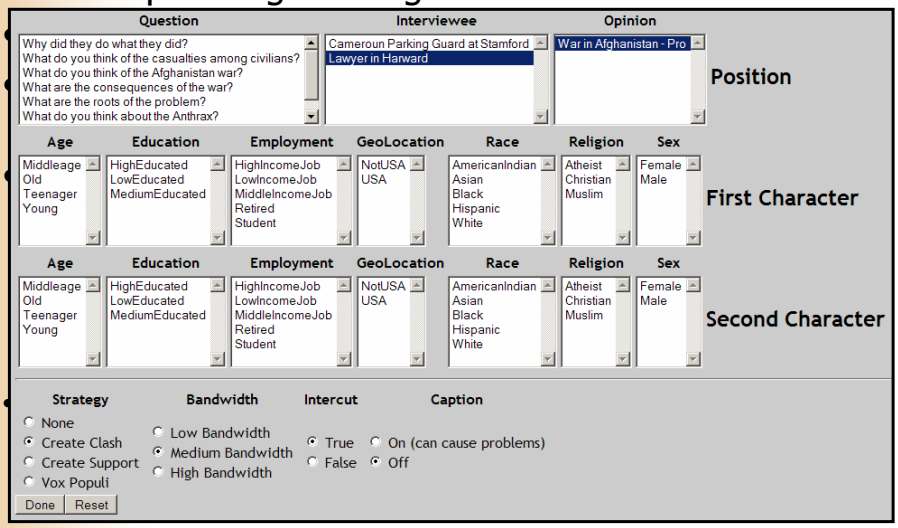


- Requires understanding of how

124

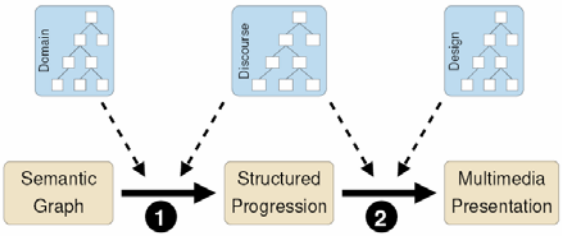
Dynamic Discourse

- Vox Populi: Argument generation in video



Role of structured progression

- Mono-media cases (text or video sequences):
  - ordering for the fragments and present them
- Presentations that use combinations of media,
  - Semantics of domain and discourse need translation to *hierarchical structures* that can be expressed through *layout*, *navigational links* or *temporal info*.
- Intermediate format is required:



126

Conclusions

- From projects described we have learned:
  - distinguish stages in process
  - separate discourse knowledge
    - Fixed and Dynamic discourse
  - mappings between domain & discourse knowledge
- Scientific challenges remain:
  - Making (MM) discourse and design knowledge explicit
  - Expressing re-usable semantics of media assets
  - Architectures for multimedia presentation generation

127

Acknowledgments

- This research was funded by:
  - the European Commission under contract FP6-027026, *Knowledge Space of semantic inference for automatic annotation and retrieval of multimedia content - K-Space*
  - the European *ITEA Passepartout project*,
  - Dutch *BSIK MultimediaN e-Culture project*

128

Literature tips

- *Roger T. Pédauque*
  Document: Form, Sign and Medium, As Reformulated for Electronic Documents
  http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/05/94/index_fr.html
- *Jacco van Ossenbruggen*
  Web Document Engineering, Part I of PhD thesis (2001):
  http://www.cwi.nl/~jrvosse/thesis/thesis-linked.pdf
- *Lloyd Rutledge, Jacco van Ossenbruggen, Lynda Hardman, and Dick C.A. Bulterman*
  Structural Distinctions Between Hypermedia Storage and Presentation
  ACM Multimedia pp. 145-150, ACM Press, Nov 1998
  http://www.cwi.nl/~lloyd/Papers/ACMMM98/acmmm98.html
- *Jane Hunter and Carl Lagoze*
  Combining RDF and XML Schemas to Enhance Interoperability Between Metadata Application Profiles
  WWW 10, Hong Kong pp. 457-466, ACM Press, May 1-5, 2001
  http://www10.org/cdrom/papers/572/
- *John Bateman, Jörg Kleinz, Thomas Kamps, and Klaus Reichenberger*
  Towards Constructive Text, Diagram, and Layout Generation for Information Presentation
  Computational Linguistics 27(3), pp. 409-449, September 2001
  http://acl.ldc.upenn.edu/J/J01/J01-3004.pdf