Semantic Web

Jacco van Ossenbruggen, Lynda Hardman
Multimedia and Human-Computer Interaction
CWI Amsterdam

With contributions from
Frank van Harmelen
AI Department
Vrije Universiteit Amsterdam

Talk overview

- **Three generations of the Web**
  - Problems with the current 2nd generation
  - The Semantic Web: a vision of the 3rd generation
- Semantic Web technology
  - XML, RDF, OWL
- 2nd generation multimedia
  - Cuypers
- Towards 3rd generation multimedia

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)

3

SW application areas

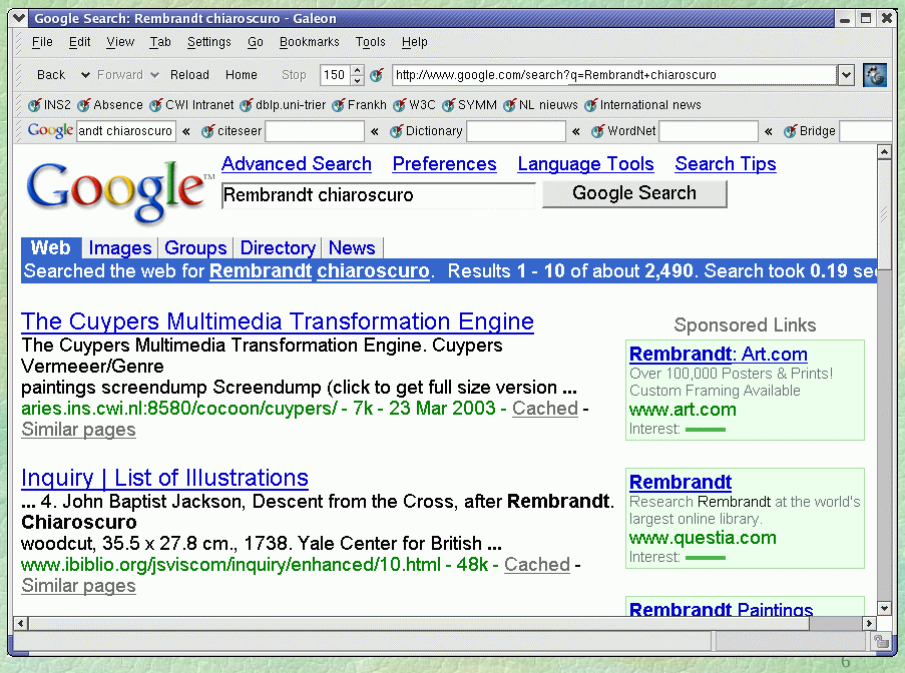
- Search engines
- Browsing on-line stores (B2C)
- Service description and integration (B2B)
- Tailored multimedia information

4

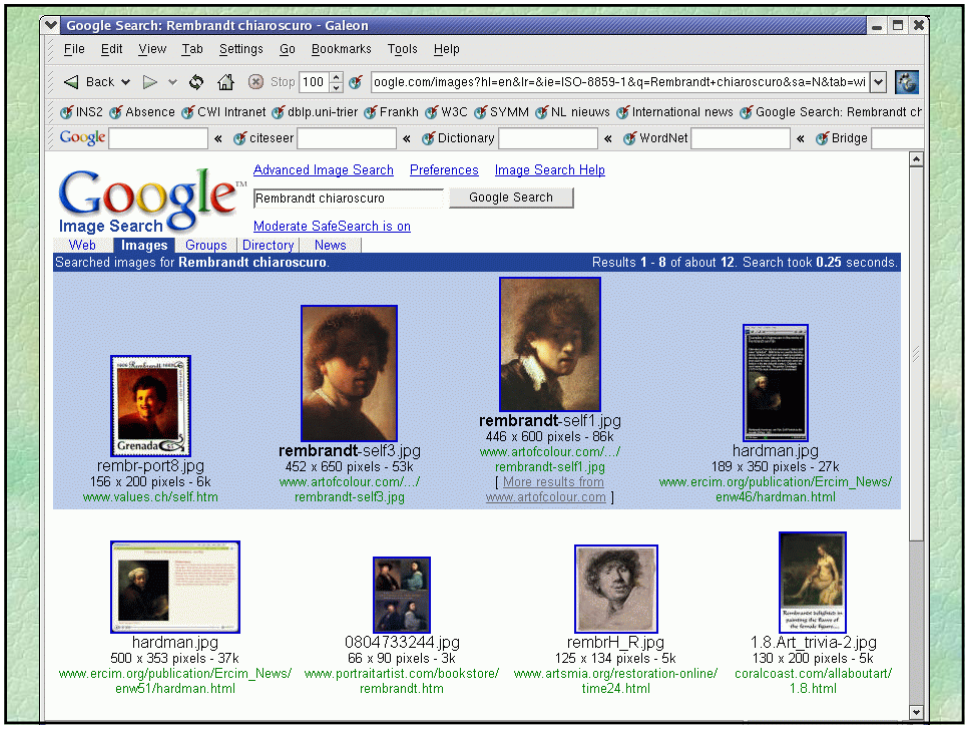
Problems with current search engines

- Current search engines = keywords:
  - high recall, low precision
  - sensitive to vocabulary
  - insensitive to implicit content

5



6

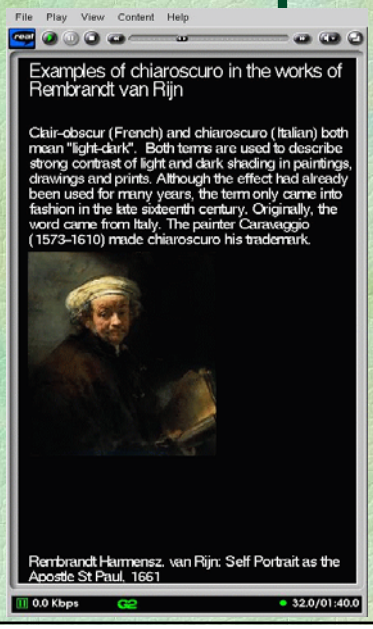


Search engines on the Semantic Web

- concept search instead of keyword search
- semantic narrowing/widening of queries
- query-answering over >1 document
- document transformation operators

8

Tailored presentation



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

9

Accessing information on the 2nd generation Web

- Students have access to material on the Web
  - Search problem
- Material is designed for "typical" student
  - No student is typical
- Some adaptivity is possible
  - Links revealed once material has been covered
- Student's knowledge level is implicit

10

Accessing information on the Semantic Web

- Students would be able to find suitable courses
- Material can be tailored for the individual
- Material can be re-used
- Models can be made of
  - The domain
  - Learner profile
  - Learning strategies
- Student's knowledge level can be made explicit
  - in terms of the domain model
  - in terms of the learning strategy

11

Talk Overview

- Three generations of the Web
  - Problems with the current 2nd generation
  - The Semantic Web: a vision of the 3rd generation
- **Semantic Web technology**
  - XML, RDF
  - DAML+OIL tutorial
- 2nd generation multimedia
  - Cuypers
- Towards 3rd generation multimedia

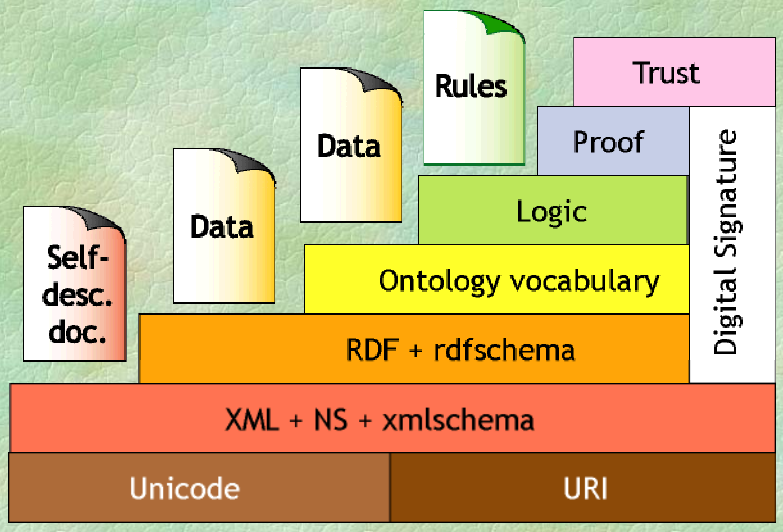
12

So what is the Semantic Web?

- The “Semantic Web Wedding Cake”
- Crash course
  - XML
  - RDF
  - RDF Schema
  - OWL

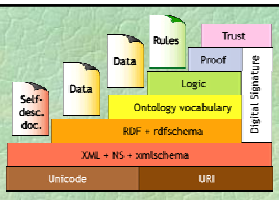
13

TBL talk at XML 2000



14

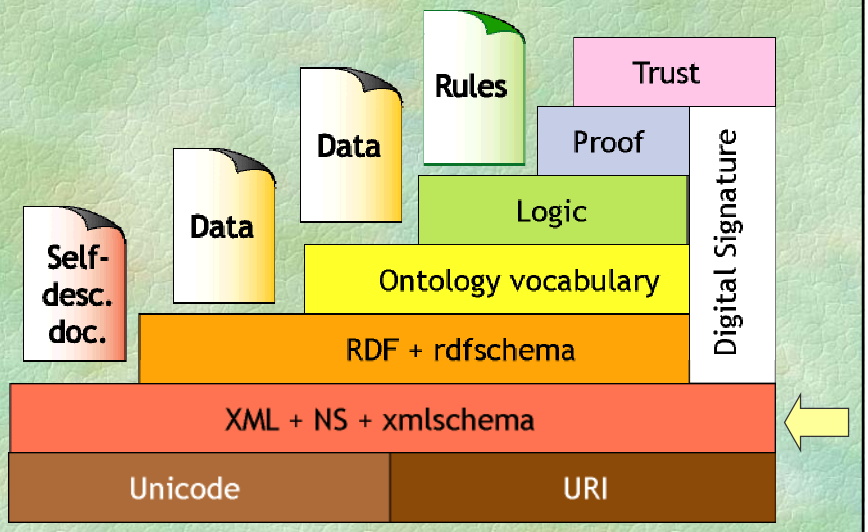
W3C Stack



- **XML**:
  - Surface syntax, no semantics
- **XML Schema**:
  - Describes structure of XML documents
- **RDF**:
  - Datamodel for “relations” between “things”
- **RDF Schema**:
  - RDF Vocabulary Definition Language
- **OWL**:
  - A more expressive Vocabulary Definition Language

15

TBL talk at XML 2000



16

XML: User definable and domain specific markup

HTML:

```
<H1>Introduction to AI</H1>
     <UL> <LI>Teacher: Frank van Harmelen
          <LI>Students: 1AI, 1I
          <LI>Requirements: none
     </UL>
```

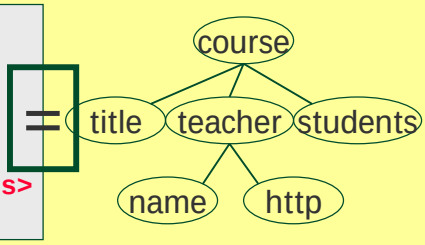
XML:

```
<course>
     <title>Introduction to AI</title>
     <teacher>Frank van Harmelen</teacher>
     <students>1AI, 1I</students>
     <req>none</req>
</course>
```

XML: document = labelled tree

- node = label + attr/values + contents

```
<course date="...">
     <title>...</title>
     <teacher>...</teacher>
          <name>...</name>
          <http>...</http>
     <students>...</students>
</course>
```



- XML Schema: grammars for describing legal trees and datatypes
- So:

why not use XML to represent semantics?

19

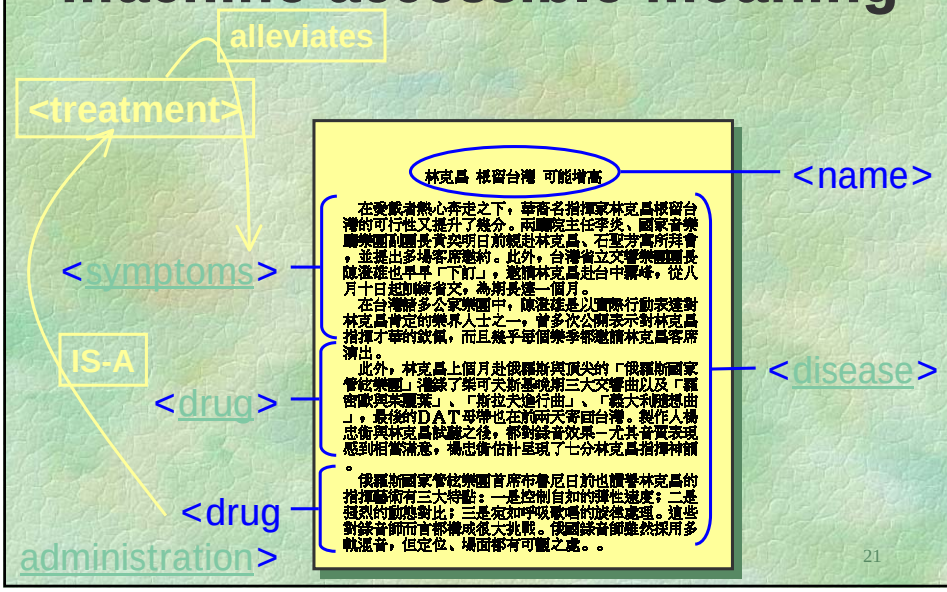
XML: limitations for semantic markup

- XML makes no commitment on:
  - ❶ Domain-specific ontological vocabulary
  - ❷ Ontological modeling primitives
- Requires pre-arranged agreement on ❶ & ❷
- Only feasible for closed collaboration
  - agents in a small & stable community
  - pages on a small & stable intranet
- Not suited for sharing Web-resources

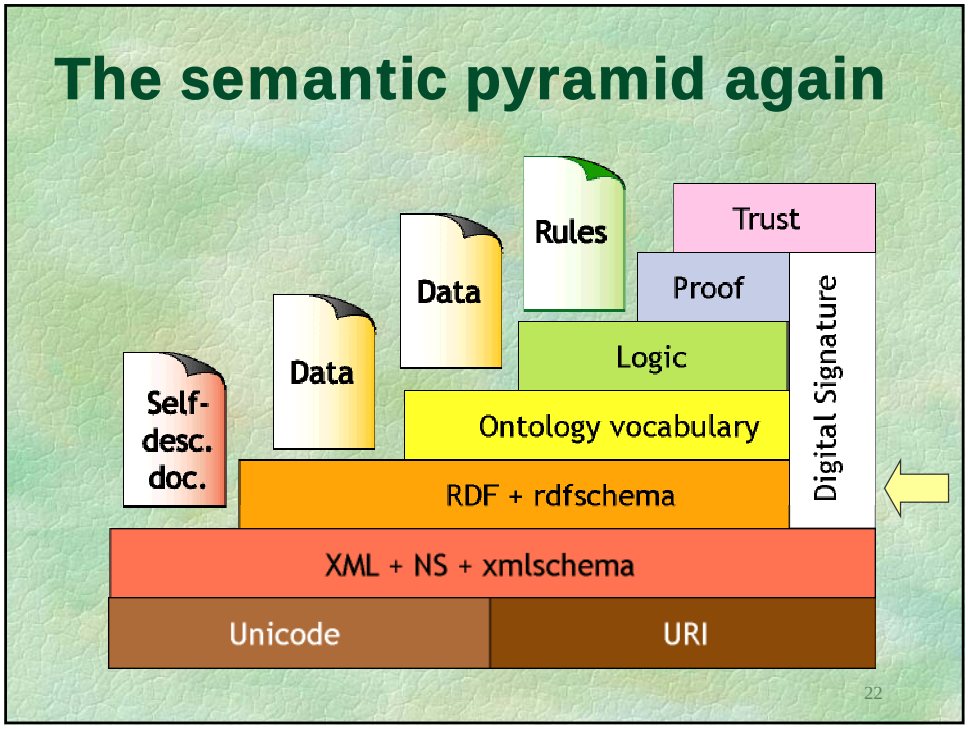
☹

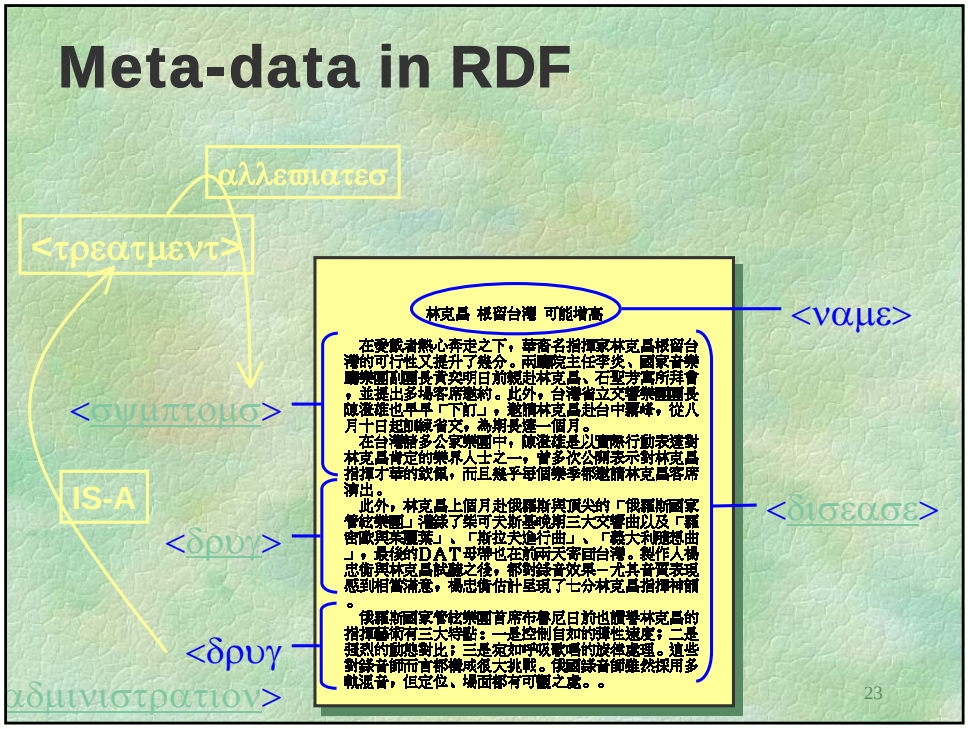
20

XML ≠ machine accessible meaning



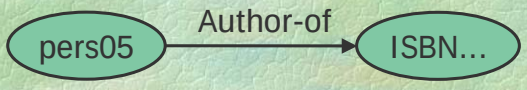
21


The semantic pyramid again
Rules
Trust
Data
Proof
Data
Logic
Self-desc. doc.
Ontology vocabulary
Digital Signature
RDF + rdfschema
XML + NS + xmlschema
Unicode
URI
22

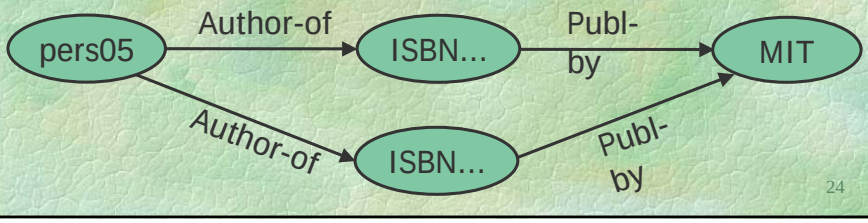

Meta-data in RDF
αλλεωιατεσ
<τρεατμεντ>
<σψμπτομσ>
IS-A
<δρυγ>
<δρυγ
αδμινιστρατιον>
林克昌 根留台灣 可能增高
<ναμε>
<δισεασε>
在愛戲者熱心奔走之下，華裔名指揮家林克昌根留台灣的可行性又提升了幾分。兩廳院主任李炎、國家音樂廳樂團副團長黃奕明日前親赴林克昌、石聖芳寓所拜會，並提出多場客席邀約。此外，台灣省立交響樂團團長陳澄雄也早早「下訂」，邀請林克昌赴台中駐峙，從八月十日起加碼省交，為期長達一個月。
在台灣諸多公家樂團中，陳澄雄是以實際行動表達對林克昌肯定的樂界人士之一，曾多次公開表示對林克昌指揮才華的欽佩，而且幾乎每個樂季都邀請林克昌客席演出。
此外，林克昌上個月赴俄羅斯與頂尖的「俄羅斯國家管絃樂團」灌錄了柴可夫斯基晚期三大交響曲以及「羅密歐與朱麗葉」、「斯拉夫進行曲」、「最大利誘惑曲」，最後的DAT母帶也在前兩天寄回台灣。製作人楊忠衡與林克昌試聽之後，都對錄音效果一尤其音質表現感到相當滿意，場忠衡估計呈現了七分林克昌指揮神韻。
俄羅斯國家管絃樂團首席布魯尼日前也讚譽林克昌的指揮藝術有三大特點：一是控制自如的彈性速度；二是強烈的動態對比；三是宛如呼吸歌唱的旋律處理。這些對錄音師而言都構成很大挑戰。俄國錄音師雖然採用多軌混音，但定位、場面都有可觀之處。。
23

RDF: graphs of triples

- Object ->Attribute-> Value triples



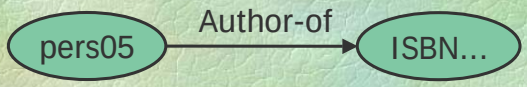
- Objects are web-resources
- Value is again an Object:
  - triples can be linked
  - data-model = graph



24

RDF is part of the Web

- Every resource has a URI
  = world-wide unique naming!



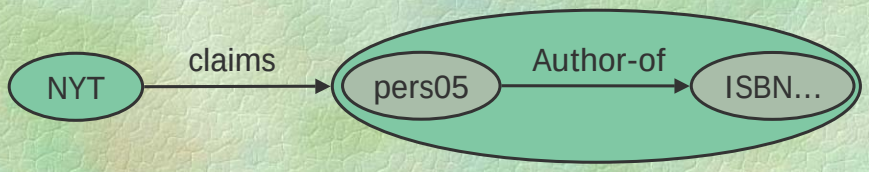
- Has XML syntax(es)

```
<rdf:Description rdf:about="#pers05">
   <authorOf>ISBN...</authorOf>
</rdf:Description>
```

25

RDF can be nested

- Any statement can be an object
  graphs can be nested - reification

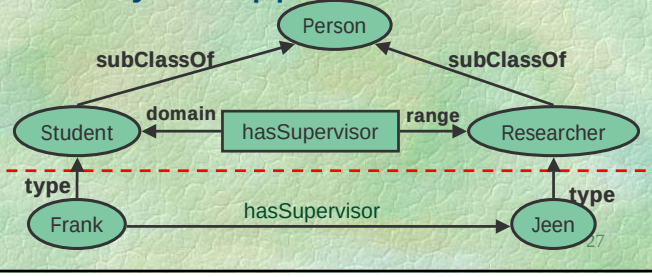



```
<rdf:Description rdf:about="#NYT">
  <claims>
    <rdf:Description rdf:about="#pers05">
      <authorOf>ISBN...</authorOf>
    </rdf:Description>
  </claims>
</rdf:Description>
```

26

What does RDF Schema add?

- Defines small **vocabulary** for RDF:
  - Class, subClassOf, type
  - Property, subPropertyOf
  - domain, range
- Vocabulary can be used to define other vocabularies for your application domain



27

RDF(S) have a (very small) formal semantics

- Defines what other statements are **implied** by a given set of RDF(S) statements
- Ensures mutual **agreement on minimal content** between parties without further contact
- In the form of “entailment rules”
- Very **simple to compute** (and not explosive in practice)

28

RDF(S) semantics: examples

- Aspirin **isOfType** Painkiller
  Painkiller **subClassOf** Drug
  ➔ Aspirin **isOfType** Drug
- aspirin alleviates headache
  alleviates **range** symptom
  ➔ headache **isOfType** symptom

29

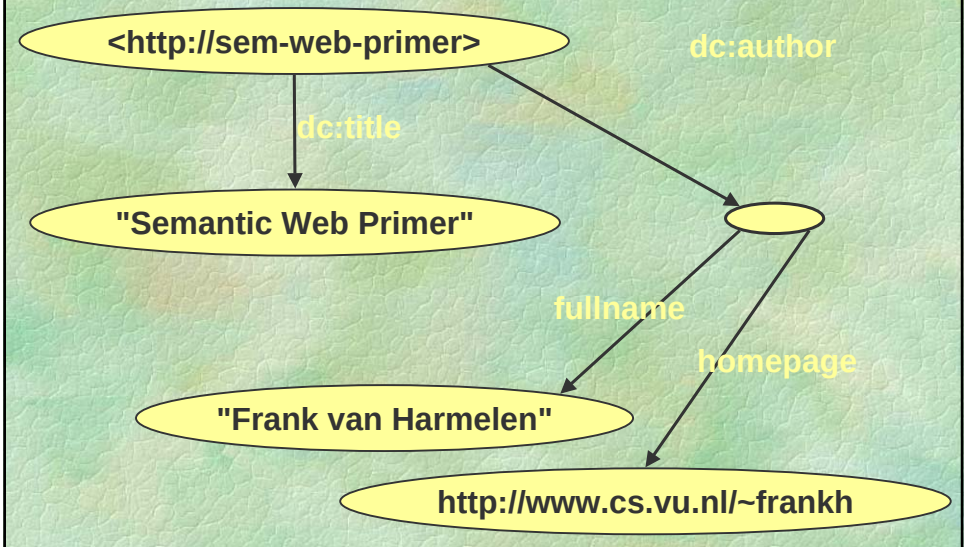
RDF(S) semantics: examples

- Ασπιριν **isOfType** Παινκιλλερ
  Παινκιλλερ **subClassOf** Δρυγ
  ➔ Ασπιριν **isOfType** Δρυγ

- ασπιριν αλλεωιατεσ ηεαδαχηε
  τρεατσ **range** σψμπτομ
  ➔ ηεαδαχηε **isOfType** σψμπτομ

30

RDF(S) syntax: graphics



31

RDF(S) syntax: XML

```
<rdf:RDF>
  <rdf:Description rdf:about="http://sem-web-primer"
   dc:title="Semantic Web Primer">
    <dc:author>
      <rdf:Description fullname="Frank van Harmelen">
        <homePage rdf:resource="http://www.cs.vu.nl/~frankh"/>
      </rdf:Description>
    </dc:author>
  </rdf:Description>
</rdf:RDF>
```

32

RDF(S) syntax: Turtle

```
<http://sem-web-primer>
  dc:title "Semantic Web Primer" ;
  dc:author [
    fullname "Frank van Harmelen";
    homePage <http://www.cs.vu.nl/~frankh>
  ] .
```

33

RDF(S) / XML relationship

- XML is a just a syntax for RDF(S)
  - (one of many)
- RDF(S) assigns meaning to some terms
  - (XML doesn't)
- This allows greater interoperability:
  - tools/tools
  - thesaurus/thesaurus
  - tools/thesaurus

34

Conclusions about RDF(S)

- Next step up from plain XML ?
  - (small) ontological commitment to modeling primitives
  - possible to define vocabulary

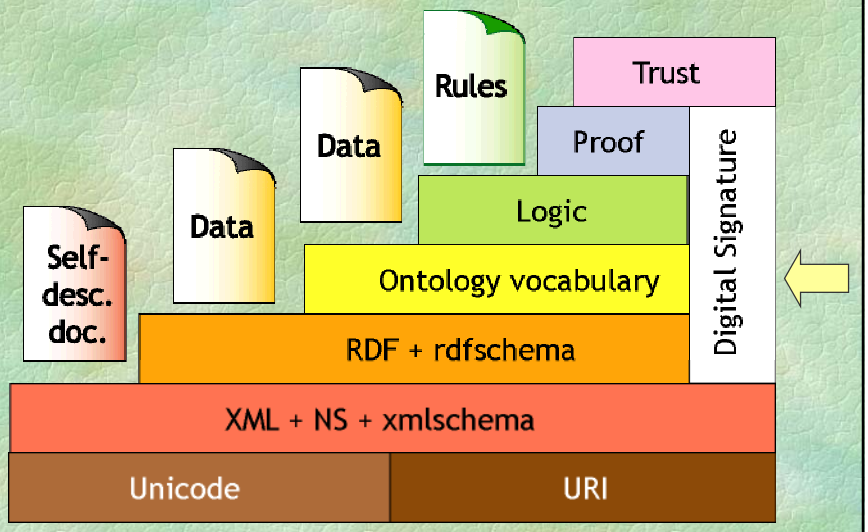
35

Things RDF(S) can't do

- equality
- enumeration
- number restrictions
  - Single-valued/multi-valued
  - Optional/required values
- inverse, symmetric, transitive
- boolean algebra
  - Union, complement
- ...

36

The semantic pyramid again



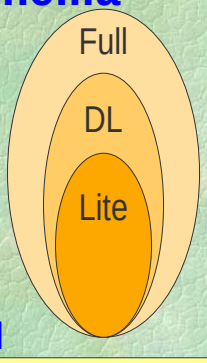
37

OWL: more expressivity

- **OWL Light**
  - (sub)classes, individuals
  - (sub)properties, domain, range
  - conjunction

  RDF Schema

  - (in)equality
  - cardinality 0/1
  - datatypes
  - inverse, transitive, symmetric
  - hasValue
  - someValuesFrom
  - allValuesFrom



- **OWL DL**
  - Negation
  - Disjunction
  - Full Cardinality
  - Enumerated types
- **OWL Full**
  - Allow meta-classes etc
  - Self-redefining power

38

OWL also has a formal semantics

- Defines what other statements are **implied** by a given set of statements
- Ensures **mutual agreement** on content (both **minimal and maximal**) between parties without further contact
- Can be used for integrity/**consistency checking**
- Hard to compute (and *rarely/sometimes/always explosive in practice*)

39

OWL semantics: minimal

- vanGogh isOfType Impressionist
  Impressionist subClassOf Painter
  ➔ **vanGogh isOfType Painter**

- vanGogh painter-of sunflowers
  painter-of domain painter
  ➔ **vanGogh isOfType painter**

40

OWL semantics: maximal

- vanGogh isOfType Impressionist
  Impressionist disjointFrom Cubist
  ➔ **NOT: vanGogh isOfType Cubist**

- painted-by has-cardinality 1
  sun-flowers painted-by vanGogh
  Picasso different-individual-from vanGogh
  ➔ **NOT: sun-flowers painted-by Picasso**

41

```
Class            Wine
  subClassOf                  PotableLiquid

  Restriction
    onProperty                #madeFromGrape
    minCardinality  1

  Restriction
    onProperty                hasMaker
    allValuesFrom               #Winery

equivalentClass               Vin
```

42

Talk overview

- Three generations of the Web
  - Problems with the current 2nd generation
  - The Semantic Web: a vision of the 3rd generation
- Semantic Web technology
  - XML, RDF, OWL
- **2nd generation multimedia**
  - **Cuypers**
- Towards 3rd generation multimedia

43

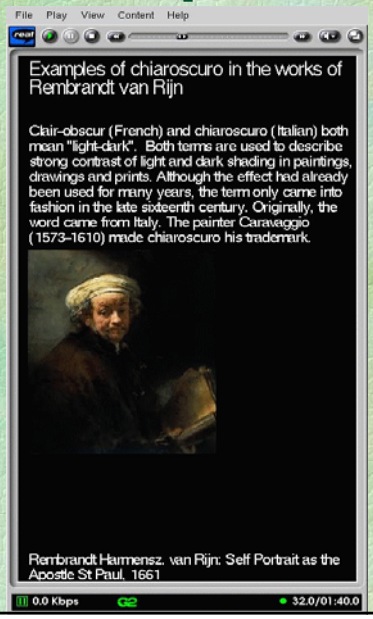
Multimedia on the Web

- Real multimedia Web content is still rare
  - Mostly bells & whistles to enhance HTML text ...
  - ... or mono-media AV-streams
- Virtually all presentations are hand-authored
  - proprietary formats that are hard to generate
  - limited support for dynamic content and multichanneling
  - most Web technology is text/page-oriented ...
  - ... with SMIL as one of the few exceptions
- Conclusion:

  Multimedia has hardly caught up with the 1st generation Web!

44

Example scenario



- User is interested in Rembrandt and wants to know about about the “chiaroscuro” technique
- System responds with textual explanation of the technique and a number of example images of its application in Rembrandt’s paintings

45

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**

46

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

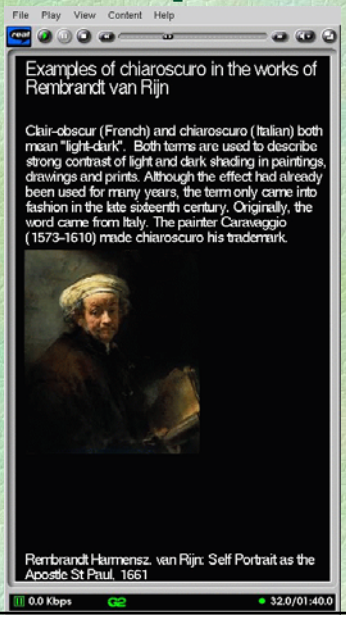
47

Talk overview

- Three generations of the Web
  - Problems with the current 2nd generation
  - The Semantic Web: a vision of the 3rd generation
- Semantic Web technology
  - XML, RDF, OWL
- 2nd generation multimedia
  - Cuypers
- **Towards 3rd generation multimedia**

50

Example - Annotated SMIL



51

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>
```

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>
    . . .
```

53

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>
```

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>
```