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

Heiner Stuckenschmidt
Intelligent Systems Group
Center for Computing Technologies
University of Bremen

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

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)
- E-learning

4

Problems with current search engines

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

5

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

6

Problems with 2nd generation on-line stores (B2C)

- manual browsing is time-consuming and inefficient
- every shopbot requires a series of wrappers
  - work only partially
  - extract only explicit information
  - must be updated frequently

7

B2C on the Semantic Web

- Software agents “understand” product descriptions
  - enabling automatic browsing
- Procedural wrapper-coding becomes declarative ontology-mapping
  - improving robustness and simplifying maintenance

8

Current B2B problems

- EDIFACT

EDIFACT S93A Sample Document

PURCHASE ORDER

```
UNB+UNOB:1+003897733:01:MFGB-PO+PARTNER ID:ZZ+000101:1050
+00000000000916++ORDERS'
UNH+1+ORDERS:S:93A:UN'
BGM+221+P1M24987E+9'
DTM+4:20000101:102'
FTX+PUR+3++PURCHASE ORDER BEFORE LINE ITEM INSTRUCTIONS'
RFF+CT:123-456'
NAD+SE+10025392::92++SUPPLIER NAME'
CTA+SR+:STEVE'
NAD+BT+B2::92++COMPAQ COMPUTER CORPORATION+P O BOX 692000
+HOUSTON+TX+77692000+US'
NAD+BY+MFUS::92++COMPAQ COMPUTER CORPORATION'
CTA+PD+:CLARETTA STRICKLAND-FULTON'
NAD+ST+CM6::92++COMPAQ COMPUTER CORPORATION+CCM6 RECEIVING DOCK:20555 SH
249+HOUSTON+TX+77070+US'
TAX+9++++++3-00105-5135-3'
CUX+2:USD:9'
PAT+1++1:1:D:45'
PCD+12:2'
TDT+20++++:::AIRBORNE'
LOC+16+COMPAQ DOCK'
TOD+2+NS+:::ORIGIN COLLECT'
IA+1+AA:EC+123456:VP'
IMD+F+8+:::PART DESCRIPTION INFORMATION
```

9

B2B on the Semantic Web

- Semantic Web technology is
  - Cheaper
  - Flexible
  - Integrated with “document” Web
- Provides interoperable semantics for
  - vertical markets (verticalnet.com)
  - horizontal markets

VerticalNet

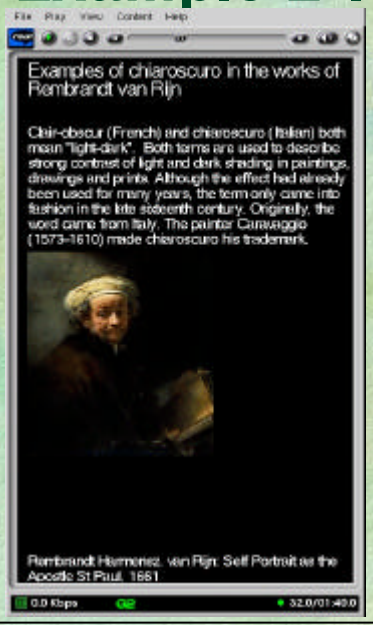
OASIS

CommerceNet

RosettaNet

10

Example E-learning 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

11

E-learning on 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

12

E-learning 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

13

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

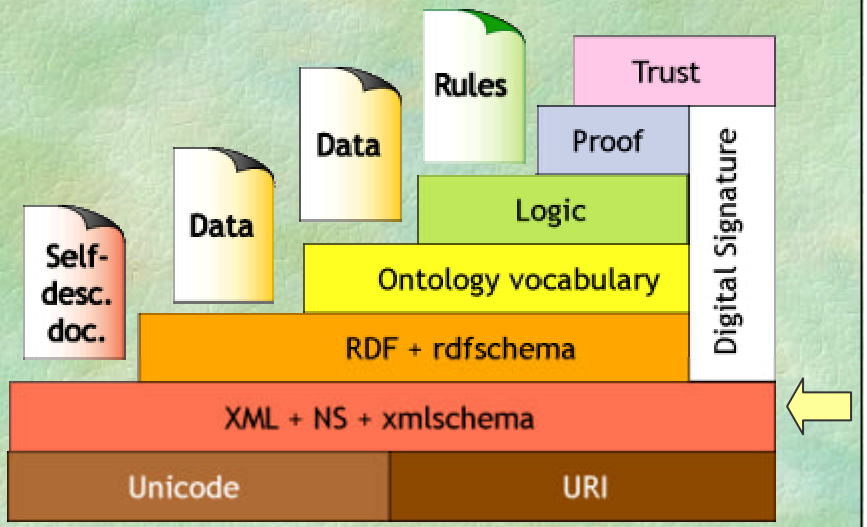
14

So what is the Semantic Web?

- Semantic Web tutorial
- Crash course XML, RDF, RDF Schema
  - To be continued next week with DAML+OIL
- The “Semantic Web Wedding Cake”

15

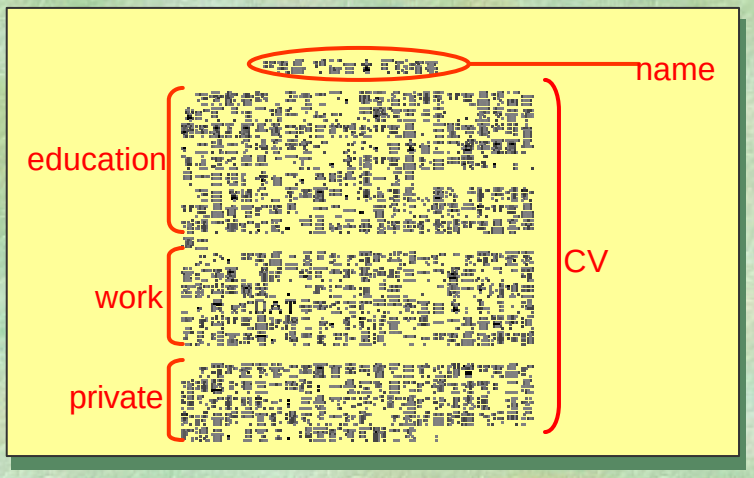
TBL talk at XML 2000



16

Machine accessible meaning

(What it's like to be a machine)



17

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>
```

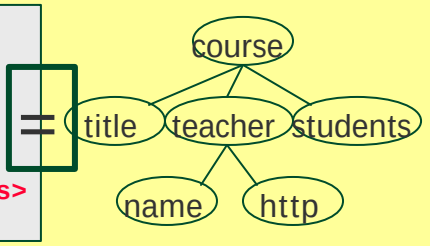
XML:

```
<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?

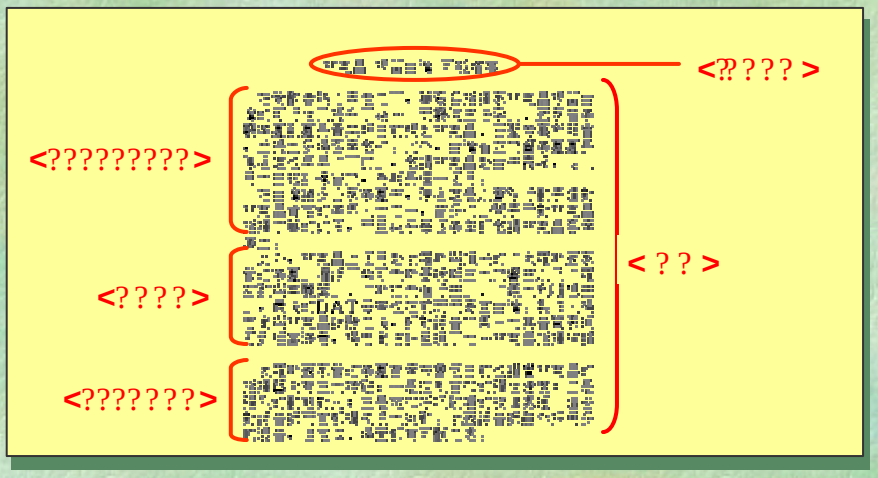
19

XML: limitations for semantic markup

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

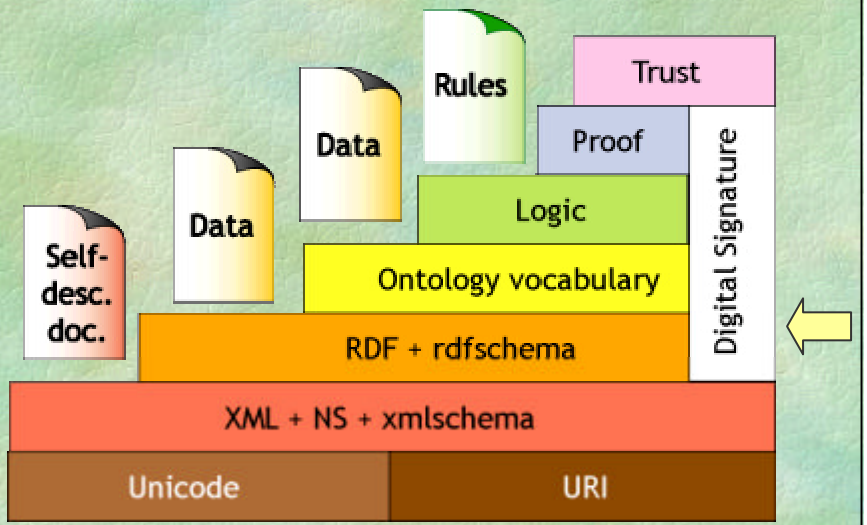
20

XML ?
machine accessible meaning



21

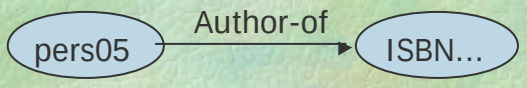
The semantic pyramid again



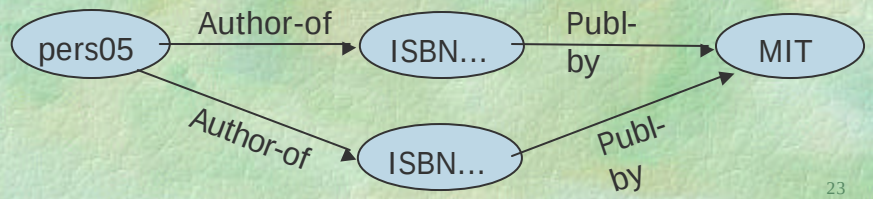
22

RDF: graphs of triples

- Object ->Attribute-> Value triples



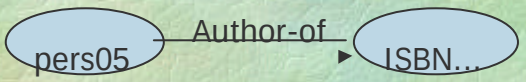
- Objects are web-resources
- Value is again an Object:
  - triples can be linked
  - data-model = graph



23

RDF is part of the Web

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



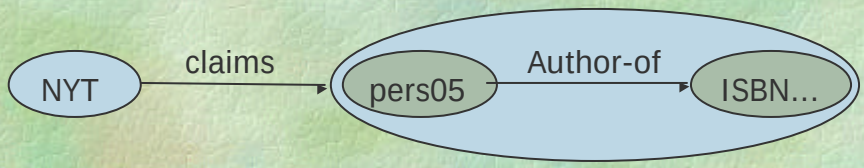
- Has XML syntax(es)

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

24

RDF can be nested

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

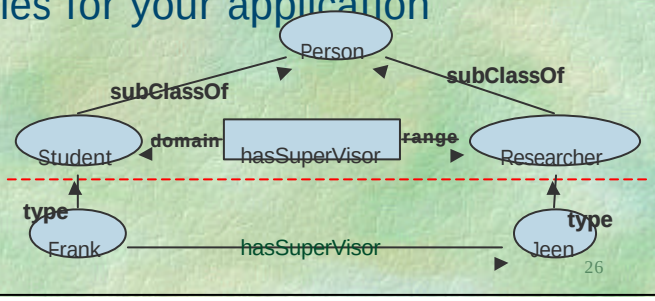



```
<rdf:Description rdf:about="#NYT">
  <claims>
    <rdf:Description rdf:about="#pers05">
      <authorOf>ISBN...</authorOf>
    </rdf:Description>
  </claims>
</rdf:Description>
```

25

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



26

RDF Schema syntax in XML

```
<rdf:Description ID="MotorVehicle">
   <rdf:type resource="http://www.w3.org/...#Class"/>
   <rdfs:subClassOf rdf:resource="http://www.w3.org/...#Resource"/>
  </rdf:Description>

<rdf:Description ID="Truck">
   <rdf:type resource="http://www.w3.org/...#Class"/>
   <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
</rdf:Description>

<rdf:Description ID="registeredTo">
   <rdf:type resource="http://www.w3.org/...#Property"/>
   <rdfs:domain rdf:resource="#MotorVehicle"/>
   <rdfs:range rdf:resource="#Person"/>
</rdf:Description>

<rdf:Description ID="ownedBy">
   <rdf:type resource="http://www.w3.org/...#Property"/>
   <rdfs:subPropertyOf rdf:resource="#registeredTo"/>
</rdf:Description>
```

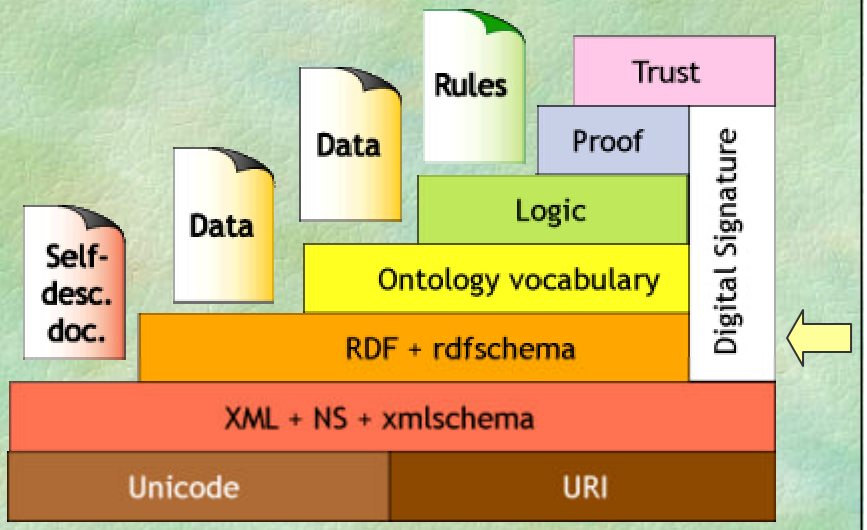
27

Conclusions about RDF(S)

- Next step up from plain XML ?
  - (small) ontological commitment to modeling primitives
  - possible to define vocabulary
- However:
  - no precisely described meaning
  - no inference model

28

Will RDF work in practice?



29

RDF vs XML

- Still many open issues in RDF
  - Serialization syntax in XML needs clean-up
  - Lack of formal semantics
- RDF Schema is not yet finished
  - Due to RDF open issues
  - XML Schema Datatypes integration
- Being RDF compatible with the rest of world is only useful when there are many other RDF applications!

30

Summary Part I

- Three generations Web
- Problems with 2nd generation Web
- Languages for 3rd generation Web
  - XML
  - RDF(S)
- Do we need RDF, or is XML sufficient?

31