


# Semantic Web Inleiding

B.J. Wielinga

(met sheets van J. Hendler, N. Noy, F. van Harmelen  
en M. Sabou)


30/10/2003 semantic web 1 1



# Cursus informatie

- Docenten: Jansen, Jansweijer, Wielinga
- Info: <http://www.swi.psy.uva.nl/semanticweb/>
- Practicum: maandag 13-16 uur, A9.01
- Elk practicum wordt afgesloten met een document dat de donderdag daarop moet zijn ingeleverd. Deze documenten worden beoordeeld. U mag in groepen van 2 werken.
- Schriftelijk tentamen medio december, geen open boek, maar wel aantekeningen
- Eindcijfer: 60% tentamen, 40% practica


30/10/2003 semantic web 1 2



# Het (huidige) WWW

- Enorme informatie bron voor velen
  - Enorme hoeveelheid data (>3 miljard paginas)
  - Honderden miljoenen gebruikers
- Zeer krachtige zoekmachines (Google)
- Maar:
  - Toch is het vaak moeilijk om de juiste informatie te vinden, te veel hits, te veel irrelevante hits
  - Non-textuele informatie is slecht geïndexeerd
  - Relaties tussen documenten worden nauwelijks gelegd => geen doorzichtige structuur
  - Er kunnen niet echt vragen gesteld worden

30/10/2003 semantic web 1 3




# Tim Berners Lee: the Semantic Web Vision

“In communicating between people using the Web, computers and networks have as their job to enable the information space, and otherwise get out of the way. But doesn't it make sense to also bring computers more into the action? – part two of the dream!

“... This creates what I call a **Semantic Web** – a web of data that can be processed directly or indirectly by machines.”

<<Weaving the Web>> pp 177.


30/10/2003 semantic web 1 4



# Why?

- To enable focussed retrieval of information from the web
- To enable machines to filter information
- To enable retrieval of multi-media information
- To link information resources
- To answer questions

30/10/2003 semantic web 1 5



# Example Question

- Where can I buy the best Chablis for a price below 30 Euro?
  - What is a Chablis?
  - What is the best Chablis?
  - Vintage?
  - Where do I find Chablis providers and their prices?
  - Do I buy a single bottle or en gros?

30/10/2003 semantic web 1 6



## The Layer Cake [TBL,XML2000]

## URI

- Unique Resource Identifier
- Cf. URL, maar er is niet noodzakelijkerwijze een document mee verbonden
- Voorbeeld: <http://www.w3.org/2001/sw/WebOnt/guide-src/wine#wijn>
- Alles tm # is de **namespace**
- Namespaces worden afgekort: vin:wijn

30/10/2003 semantic web 1 14

## URI

- Elk object in de wereld (concreet of abstract) heeft een unieke naam: Unique Resource Identifier
- De URI's worden gebruikt in beschrijvingen van objecten in de wereld.

30/10/2003 semantic web 1 15

## What is XML?

- **XML: eXtensible Markup Language**
- **Subset of SGML, but more efficient and lightweight**
- **XML:**
  - **Extensible:** tags can be defined; can be extended to lots of different applications.
  - **Markup language:** language used to mark up data.
  - **Meta Language:** Language used to represent other languages.

30/10/2003 semantic web 1 16

## XML v. HTML

- **contrast with HTML: XML is Extensible:**
  - **HTML:** restricted set of tags, e.g. <TABLE>, <H1>, <B>, etc.
  - **XML:** you can create your own tags
- **Example: Put a library catalog on the web.**
  - **HTML:** You are stuck with regular HTML tags, e.g. H1, H3, etc.
  - **XML:** You can create your own set of tags: TITLE, AUTHOR, DATE, PUBLISHER, etc.

30/10/2003 semantic web 1 17

## Book Catalog in HTML

```
<HTML>
<BODY>
<H1>Harry Potter</H1>
<H2>J. K. Rowling</H2>
<H3>1999</H3>
<H3>Scholastic</H3>
</BODY>
</HTML>
```

HTML conveys the "look and feel!" of your page.

As a human, it is easy to pick out the publisher.

But, how would a computer pick out the publisher?

Answer: XML

30/10/2003 semantic web 1

**Book Catalog in XML**

```

<BOOK>
  <TITLE>Harry Potter</TITLE>
  <AUTHOR>J. K. Rowling</AUTHOR>
  <DATE>1999</DATE>
  <PUBLISHER>Scholastic</PUBLISHER>
</BOOK>
  
```

Look at the new tags!  
A Human and a computer can now easily extract the publisher data.

30/10/2003 semantic web 1 19

But, XML ≠ machine accessible meaning

To your machine, the tags look like this....

30/10/2003 semantic web 1 20

**Kennisrepresentatie in het SW**

- XML biedt structuur maar geen kennis representatie
- RDF is een heel eenvoudig model voor kennisrepresentatie, vergelijkbaar met een psychologisch model van het geheugen uit eind zestiger jaren van de 20e eeuw (N. Frijda)
- A-R-B (triple)

30/10/2003 semantic web 1 21

**RDF (Resource Description Framework)**

- RDF provides a way of describing resources via metadata (data about data) It restricts the description of resources to **triples (subject, predicate, object)**
- It provides interoperability between applications that exchange machine understandable information on the Web.
- The original broad goal of RDF was to define a mechanism for describing resources that makes no assumptions about a particular application domain, nor defines (a priori) the semantics of any application domain.
- Uses XML as the interchange syntax.
- Provides a **lightweight** ontology system.

The formal specification of RDF is available at:  
<http://www.w3.org/TR/REC-rdf-syntax/>

30/10/2003 semantic web 1 22

**RDF schema (RDFS)**


- Adds constructs to RDF:
  - Concepts with properties
  - SubclassOf (class hierarchy)
  - Properties
  - SubpropertyOf

30/10/2003 semantic web 1 23

**RDF(S) Terminology and Semantics**


- **Classes and a class hierarchy**
  - All classes are instances of `rdfs:Class`
  - A class hierarchy is defined by `rdfs:subClassOf`
- **Instances (individuals) of a class**
  - Defined by `rdf:type`
- **Properties**
  - Properties are global:  
A property name in one place is the same as the property name in another (assuming the same namespace)
  - Properties form a hierarchy, too (`rdfs:subPropertyOf`)

30/10/2003 semantic web 1 24

 **Domain and Range of property**

- Domain of a property – the class (or classes) that have the property
  - More precisely: class (or classes) instances of which can have the property
- Range of a property – the class (or classes) to which property values belong

30/10/2003 semantic web 1 25

 **Domain and Range (2)**


- Range can be a class or a literal
- Consider the *colour* property
  - Domain: Wine
  - Range: literal (string) or
  - Range: class colour

30/10/2003 semantic web 1 26

 **KR uitbreidingen van RDF(S)**


- DAML+OIL
- OWL (W3C candidate standard)
  - Meer eigenschappen (bv. inverseOf)
  - Cardinaliteit
  - Logische expressies (geen FOL)
  - Descriptions

30/10/2003 semantic web 1 27

 **SW Languages**


- The languages differ in their
  - syntax
    - We are not concerned with it here – The essence is the conceptual representation
  - terminology
    - Class-concept
    - Instance-object
    - property-property-slot
  - expressivity
    - What we can express in some languages, we cannot express in others
  - semantics
    - The same statements may mean different things in different languages
  - tools

30/10/2003 semantic web 1 28

 **Ontological commitments**

- Wat voor soort dingen bestaan er in de externe wereld?
- HTML: documents (href)
- RDF:
  - objects (Resources)
  - binary relations (triples)
- OWL: descriptions

30/10/2003 semantic web 1 29

 **What Is An Ontology**

- An ontology is an explicit description of a domain:
  - concepts
  - properties and attributes of concepts
  - constraints on properties and attributes
  - Individuals (*often, but not always*)
- An ontology defines
  - a common vocabulary
  - a shared understanding

30/10/2003 semantic web 1 30

## Ontology Examples

- Taxonomies on the Web
  - Yahoo! categories
- Catalogs for on-line shopping
  - Amazon.com product catalog
- Domain-specific standard terminology
  - Unified Medical Language System (UMLS)
  - UNSPSC - terminology for products and services
- General ontologies: Wordnet, Cyc, ...

30/10/2003 semantic web 1 31

30/10/2003 semantic web 1 32

## Opdracht 1

- Bouw een ontologie voor wijn
- Tool: *Triple20* (vgl *Protégé*)
- Stappen:
  - Definieer concepten
  - Breng structuur aan (klasse hiërarchie)
  - Definieer properties
  - Definieer instanties

30/10/2003 semantic web 1 33

## Wines and Wineries

30/10/2003 semantic web 1 34

## Levels in the Hierarchy

30/10/2003 35