

Chapter 1

The Semantic Web Vision

Grigoris Antoniou
Frank van Harmelen

adapted and shortened
by Gerd Stumme

1

Chapter 1

A Semantic Web Primer

Lecture Outline

1. Today's Web
2. The Semantic Web Impact
3. Semantic Web Technologies
4. A Layered Approach

2

Chapter 1

A Semantic Web Primer

Today's Web

- Most of today's Web content is suitable for **human consumption**
 - Even Web content that is generated automatically from databases is usually presented without the original structural information found in databases
- Typical Web uses today people's
 - seeking and making use of information, searching for and getting in touch with other people, reviewing catalogs of online stores and ordering products by filling out forms

3

Chapter 1

A Semantic Web Primer

Keyword-Based Search Engines

- Current Web activities are not particularly well supported by software tools
 - Except for **keyword-based search engines** (e.g. Google, AltaVista, Yahoo)
- The Web would not have been the huge success it was, were it not for search engines

4

Chapter 1

A Semantic Web Primer

Problems of Keyword-Based Search Engines

- High recall, low precision.
- Low or no recall
- Results are highly sensitive to vocabulary
- Results are single Web pages
- Human involvement is necessary to interpret and combine results
- Results of Web searches are not readily accessible by other software tools

5

Chapter 1

A Semantic Web Primer

The Semantic Web Approach

- Represent Web content in a form that is more easily machine-processable.
- Use intelligent techniques to take advantage of these representations.
- The Semantic Web will gradually evolve out of the existing Web, it is not a competition to the current WWW

7

Chapter 1

A Semantic Web Primer

The Key Problem of Today's Web

- The meaning of Web content is not machine-accessible: **lack of semantics**
- It is simply difficult to distinguish the meaning between these two sentences:

I am a professor of computer science.

*I am a professor of computer science,
you may think. Well, . . .*

6

Chapter 1

A Semantic Web Primer

Lecture Outline

1. Today's Web
2. **The Semantic Web Impact**
3. Semantic Web Technologies
4. A Layered Approach

8

Chapter 1

A Semantic Web Primer

The Semantic Web Impact – Knowledge Management

- Knowledge management concerns itself with acquiring, accessing, and maintaining knowledge within an organization
- Key activity of large businesses: internal knowledge as an intellectual asset
- It is particularly important for international, geographically dispersed organizations
- Most information is currently available in a weakly structured form (e.g. text, audio, video)

9

Chapter 1

A Semantic Web Primer

Semantic Web Enabled Knowledge Management

- Knowledge will be organized in conceptual spaces according to its meaning.
- Automated tools for maintenance and knowledge discovery
- Semantic query answering
- Query answering over several documents
- Defining who may view certain parts of information (even parts of documents) will be possible.

11

Chapter 1

A Semantic Web Primer

Limitations of Current Knowledge Management Technologies

- *Searching information*
 - Keyword-based search engines
- *Extracting information*
 - human involvement necessary for browsing, retrieving, interpreting, combining
- *Maintaining information*
 - inconsistencies in terminology, outdated information.
- *Viewing information*
 - Impossible to define views on Web knowledge

10

Chapter 1

A Semantic Web Primer

The Semantic Web Impact – B2C Electronic Commerce

- A typical scenario: user visits one or several online shops, browses their offers, selects and orders products.
- Ideally humans would visit all, or all major online stores; but too time consuming
- **Shopbots** are a useful tool

12

Chapter 1

A Semantic Web Primer

Limitations of Shopbots

- They rely on wrappers: extensive programming required
- Wrappers need to be reprogrammed when an online store changes its outfit
- Wrappers extract information based on textual analysis
 - Error-prone
 - Limited information extracted

13

Chapter 1

A Semantic Web Primer

The Semantic Web Impact – B2B Electronic Commerce

- Greatest economic promise
- Currently relies mostly on EDI
 - Isolated technology, understood only by experts
 - Difficult to program and maintain, error-prone
 - Each B2B communication requires separate programming
- Web appears to be perfect infrastructure
 - But B2B not well supported by Web standards

15

Chapter 1

A Semantic Web Primer

Semantic Web Enabled B2C Electronic Commerce

- Software agents that can interpret the product information and the terms of service.
 - Pricing and product information, delivery and privacy policies will be interpreted and compared to the user requirements.
- Information about the reputation of shops
- Sophisticated shopping agents will be able to conduct automated negotiations

14

Chapter 1

A Semantic Web Primer

Semantic Web Enabled B2B Electronic Commerce

- Businesses enter partnerships without much overhead
- Differences in terminology will be resolved using standard abstract domain models
- Data will be interchanged using translation services.
- Auctioning, negotiations, and drafting contracts will be carried out automatically (or semi-automatically) by software agents

16

Chapter 1

A Semantic Web Primer

Lecture Outline

1. Today's Web
2. The Semantic Web Impact
3. Semantic Web Technologies
4. A Layered Approach

17

Chapter 1

A Semantic Web Primer

On HTML

- Web content is currently formatted for human readers rather than programs
- HTML is the predominant language in which Web pages are written (directly or using tools)
- Vocabulary describes presentation

19

Chapter 1

A Semantic Web Primer

Semantic Web Technologies

- Explicit Metadata
- Ontologies
- Logic and Inference
- Agents

18

Chapter 1

A Semantic Web Primer

An HTML Example

```
<h1>Agilitas Physiotherapy Centre</h1>
Welcome to the home page of the Agilitas Physiotherapy Centre. Do
you feel pain? Have you had an injury? Let our staff Lisa Davenport,
Kelly Townsend (our lovely secretary) and Steve Matthews take care
of your body and soul.
<h2>Consultation hours</h2>
Mon 11am - 7pm<br>
Tue 11am - 7pm<br>
Wed 3pm - 7pm<br>
Thu 11am - 7pm<br>
Fri 11am - 3pm<p>
But note that we do not offer consultation during the weeks of the
<a href="...">State Of Origin</a> games.
```

20

Chapter 1

A Semantic Web Primer

Problems with HTML

- Humans have no problem with this
- Machines (software agents) do:
 - How distinguish therapists from the secretary,
 - How determine exact consultation hours
 - They would have to follow the link to the State Of Origin games to find when they take place.

21

Chapter 1

A Semantic Web Primer

Explicit Metadata

- This representation is far more easily processable by machines
- **Metadata: data about data**
 - Metadata capture part of the *meaning of data*
- Semantic Web does not rely on text-based manipulation, but rather on machine-processable metadata

23

Chapter 1

A Semantic Web Primer

A Better Representation

```
<company>
  <treatmentOffered>Physiotherapy</treatmentOffered>
  <companyName>Agilitas Physiotherapy
  Centre</companyName>
  <staff>
    <therapist>Lisa Davenport</therapist>
    <therapist>Steve Matthews</therapist>
    <secretary>Kelly Townsend</secretary>
  </staff>
</company>
```

22

Chapter 1

A Semantic Web Primer

What are(n't) meta data good for?

- Consider a typical web page:

The screenshot shows the homepage for the 11th International World Wide Web Conference (WWW 2002) held in Hawaii from May 7-11, 2002. The page features a navigation menu on the left with links for Conference Proceedings, Call for Participation, Program, Registration Information, Hotel Accommodation, Conference Committee, Sponsorship/Exhibition Opportunities, Volunteer Information, Information about Hawaii, and Previous & Future WWW Conferences. The main content area includes the conference title, location (Sheraton Waikiki Hotel, Honolulu, Hawaii, USA), dates, and a list of registered participants from various countries. A 'REGISTER NOW' button is prominently displayed. Below, there is a section for 'FEATURED SPEAKERS (CONFIRMED)' listing speakers like Tim Berners-Lee and Richard A. DeMillo.

- Markup consists of:
 - rendering information (e.g., font size and colour)
 - Hyper-links to related content
- Semantic content is accessible to humans but not (easily) to computers...

24

Chapter 1

A Semantic Web Primer

What information can we see...

WWW2002
 The eleventh international world wide web conference
 Sheraton waikiki hotel
 Honolulu, hawaii, USA
 7-11 may 2002
 1 location 5 days learn interact
 Registered participants coming from
 australia, canada, chile denmark, france, germany, ghana, hong kong,
 india, ireland, italy, japan, malta, new zealand, the netherlands,
 norway, singapore, switzerland, the united kingdom, the united states,
 vietnam, zaire
 Register now
 On the 7th May Honolulu will provide the backdrop of the eleventh
 international world wide web conference. This prestigious event ...
 Speakers confirmed
 Tim berners-lee
 Tim is the well known inventor of the Web, ...
 Ian Foster
 Ian is the pioneer of the Grid, the next generation internet ...

Solution: XML markup with "meaningful" tags?

```

<name> WWW2002
</name>
<location> Sheraton waikiki hotel
</location>
<date> 7-11 may 2002
</date>
<slogan> 1 location 5 days learn interact
</slogan>
<participants> australia, canada, chile denmark, france, germany, ghana, hong kong,
india, ireland, italy, japan, malta, new zealand, the netherlands,
norway, singapore, switzerland, the united kingdom, the united states,
vietnam, zaire
</participants>
<introduction> On the 7th May Honolulu will provide the backdrop of the eleventh
international world wide web conference. This prestigious event ...
</introduction>
<speaker> Tim berners-lee
</speaker>
<bio> Tim is the well known inventor of the Web, ...
</bio>...
  
```

What information can a machine see...

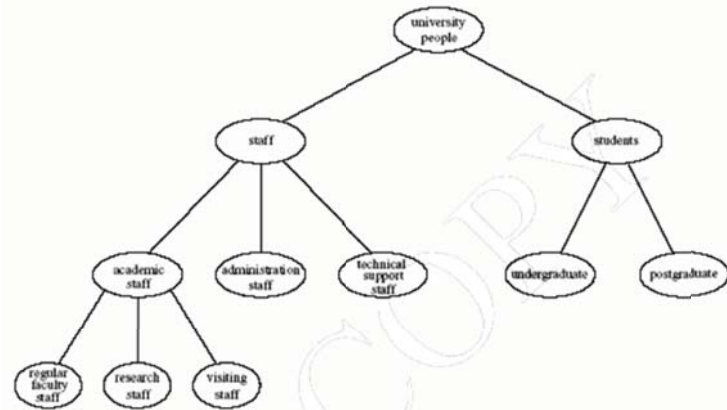
WWW2002
 The eleventh international world wide web conference
 Sheraton waikiki hotel
 Honolulu, hawaii, USA
 7-11 may 2002
 1 location 5 days learn interact
 Registered participants coming from
 australia, canada, chile denmark, france, germany, ghana, hong kong,
 india, ireland, italy, japan, malta, new zealand, the netherlands,
 norway, singapore, switzerland, the united kingdom, the united states,
 vietnam, zaire
 Register now
 On the 7th May Honolulu will provide the backdrop of the eleventh
 international world wide web conference. This prestigious event ...
 Speakers confirmed
 Tim berners-lee
 Tim is the well known inventor of the Web, ...
 Ian Foster
 Ian is the pioneer of the Grid, the next generation internet ...

But What About...

```

<conf> WWW2002
</conf>
<place> Sheraton waikiki hotel
</place>
<date> 7-11 may 2002
</date>
<slogan> 1 location 5 days learn interact
</slogan>
<participants> australia, canada, chile denmark, france, germany, ghana, hong kong,
india, ireland, italy, japan, malta, new zealand, the netherlands,
norway, singapore, switzerland, the united kingdom, the united states,
vietnam, zaire
</participants>
<introduction> On the 7th May Honolulu will provide the backdrop of the eleventh
international world wide web conference. This prestigious event ...
</introduction>
<speaker> Tim berners-lee
</speaker>
<bio> Tim is the well known inventor of the Web, ...
</bio>...
  
```


Example of a Class Hierarchy



33

Chapter 1

A Semantic Web Primer

The Role of Ontologies in Web Search

- Ontologies are useful for improving the accuracy of Web searches
 - search engines can look for pages that refer to a precise concept in an ontology
- Web searches can exploit generalization/specialization information
 - If a query fails to find any relevant documents, the search engine may suggest to the user a more general query.
 - If too many answers are retrieved, the search engine may suggest to the user some specializations.

35

Chapter 1

A Semantic Web Primer

The Role of Ontologies on the Web

- Ontologies provide a shared understanding of a domain: **semantic interoperability**
 - overcome differences in terminology
 - mappings between ontologies
- Ontologies are useful for the organization and navigation of Web sites

34

Chapter 1

A Semantic Web Primer

Web Ontology Languages

RDF Schema

- RDF is a data model for objects and relations between them
- RDF Schema is a vocabulary description language
- Describes properties and classes of RDF resources
- Provides semantics for generalization hierarchies of properties and classes

36

Chapter 1

A Semantic Web Primer

Web Ontology Languages (2)

OWL

- A richer ontology language
- relations between classes
 - e.g., disjointness
- cardinality
 - e.g. “exactly one”
- richer typing of properties
- characteristics of properties (e.g., symmetry)

37

Chapter 1

A Semantic Web Primer

An Inference Example

$\text{prof}(X) \rightarrow \text{faculty}(X)$
 $\text{faculty}(X) \rightarrow \text{staff}(X)$
 $\text{prof}(\text{michael})$

We can deduce the following **conclusions**:

$\text{faculty}(\text{michael})$
 $\text{staff}(\text{michael})$
 $\text{prof}(X) \rightarrow \text{staff}(X)$

39

Chapter 1

A Semantic Web Primer

Logic and Inference

- Logic is the discipline that studies the principles of reasoning
- Formal languages for expressing knowledge
- Well-understood formal semantics
 - Declarative knowledge: we describe what holds without caring about how it can be deduced
- Automated reasoners can deduce (infer) conclusions from the given knowledge

38

Chapter 1

A Semantic Web Primer

Logic versus Ontologies

- The previous example involves knowledge typically found in ontologies
 - Logic can be used to uncover ontological knowledge that is implicitly given
 - It can also help uncover unexpected relationships and inconsistencies
- Logic is more general than ontologies
 - It can also be used by intelligent agents for making decisions and selecting courses of action

40

Chapter 1

A Semantic Web Primer

Tradeoff between Expressive Power and Computational Complexity

- The more expressive a logic is, the more computationally expensive it becomes to draw conclusions
 - Drawing certain conclusions may become impossible if non-computability barriers are encountered.
- Our previous examples involved rules “*If conditions, then conclusion,*” and only finitely many objects
 - This subset of logic is tractable and is supported by efficient reasoning tools

41

Chapter 1

A Semantic Web Primer

Typical Explanation Procedure

- Facts will typically be traced to some Web addresses
 - The trust of the Web address will be verifiable by agents
- Rules may be a part of a shared commerce ontology or the policy of the online shop

43

Chapter 1

A Semantic Web Primer

Inference and Explanations

- Explanations: the series of inference steps can be retraced
- They increase users' confidence in Semantic Web agents
- Activities between agents: create or validate proofs

42

Chapter 1

A Semantic Web Primer

Software Agents

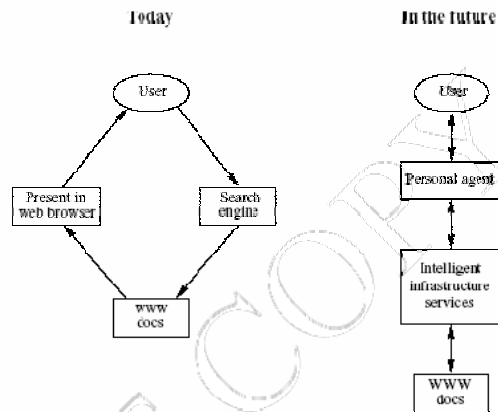
- Software agents work autonomously and proactively
 - They evolved out of object oriented and component-based programming
- A personal agent on the Semantic Web will:
 - receive some tasks and preferences from the person
 - seek information from Web sources, communicate with other agents
 - compare information about user requirements and preferences, make certain choices
 - give answers to the user

44

Chapter 1

A Semantic Web Primer

Intelligent Personal Agents



45

Chapter 1

A Semantic Web Primer

Semantic Web Agent Technologies (2)

- Further technologies (orthogonal to the Semantic Web technologies)
 - Agent communication languages
 - Formal representation of beliefs, desires, and intentions of agents
 - Creation and maintenance of user models.

47

Chapter 1

A Semantic Web Primer

Semantic Web Agent Technologies

- Metadata
 - Identify and extract information from Web sources
- Ontologies
 - Web searches, interpret retrieved information
 - Communicate with other agents
- Logic
 - Process retrieved information, draw conclusions

46

Chapter 1

A Semantic Web Primer

Lecture Outline

1. Today's Web
2. The Semantic Web Impact
3. Semantic Web Technologies
4. [A Layered Approach](#)

48

Chapter 1

A Semantic Web Primer

A Layered Approach

- The development of the Semantic Web proceeds in steps
 - Each step building a layer on top of another

Principles:

- Downward compatibility
- Upward partial understanding

49

Chapter 1

A Semantic Web Primer

Semantic Web Layers

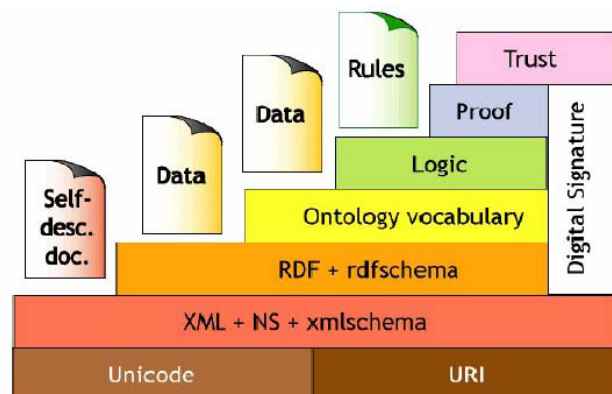
- XML layer
 - Syntactic basis
- RDF layer
 - RDF basic data model for facts
 - RDF Schema simple ontology language
- Ontology layer
 - More expressive languages than RDF Schema
 - Current Web standard: OWL

51

Chapter 1

A Semantic Web Primer

The Semantic Web Layer Tower



50

Chapter 1

A Semantic Web Primer

Semantic Web Layers (2)

- Logic layer
 - enhance ontology languages further
 - application-specific declarative knowledge
- Proof layer
 - Proof generation, exchange, validation
- Trust layer
 - Digital signatures
 - recommendations, rating agencies

52

Chapter 1

A Semantic Web Primer

Book Outline

2. Structured Web Documents in XML
3. Describing Web Resources in RDF
4. Web Ontology Language: OWL
5. Logic and Inference: Rules
6. Applications
7. Ontology Engineering
8. Conclusion and Outlook