Introduction to the semantic web
or
An overview of tomorrow's web...

Pascal Mainini

`pascal <at> impressionet <dot> ch`

`gurke <at> netlabs <dot> org`

2009-06-06

About me

- I'm Pascal Mainini
- Open minded, critical hacker
- Started with computers almost 20 years ago, working professionally since nearly 10 years
- Currently working as a network and security specialist
- *Note: I don't like beeing photographed or recorded otherwise, thanks!*

About this speech

This speech will

- Give an idea of what the semantic web is about
- Give an overview of underlying technologies
- Serve as a starting point for further explorations of the semantic web

About this speech

This speech will **NOT**

- Provide an exact mathematical background (I'm too lame for that...)
- Give an in-depth tutorial of the technologies used (URIs, XML, RDF...)
- Allow you to start working with semweb-technologies without investing any further work

Note: If you have questions - ask them right away...!

About this speech

The slides and additional material can (soon) be found at:

`http://impressionet.ch/semwebspeech3`

Problems of the current web

- A gigantic bunch of information, a large diversity of formats
- This information is stored in a form understandable for humans (which is great!)
- It's not that easy for a machine to understand (which is not so great...). . .
- **Thus, information is hard to find and reuse**

Solution approaches

- Improve the usage of what's already there
  - Better search-techniques, artificial intelligence...
  - *Hard to accomplish, results not satisfying*
- Provide the information in a way better understandable by machines
  - This requires standardised formats
  - These must be formally correct, simple and easily extensible

The semantic web

This is where the idea of the semantic web comes in.

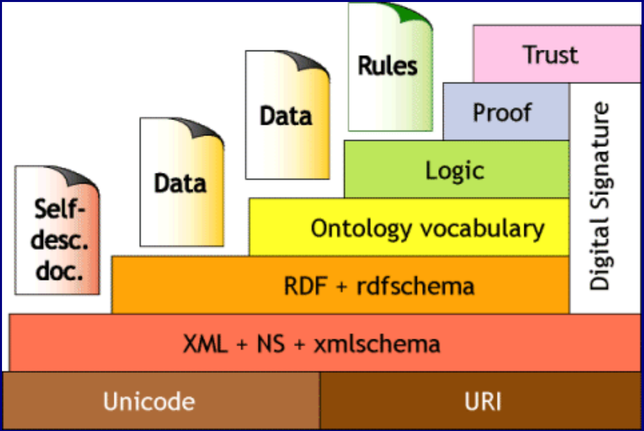
A full set of standards to accomplish this has been created by the W3C.

Technologies used base (not exclusively) on **XML** and **URI**s

On top of these follow **RDF**, **rdfschema** and **OWL**.

This is called the *semantic web layer cake*, let's have look at it...

The layer cake



[w3.org]

This talk focuses on layers 3 and 4

History

Important points in the history of the semantic web:

- Some initial work during 1997-1998
- In 1999
  - February: First recommendation, RDF model and syntax
  - March: rdfschema proposal
- February 2004: A suite of RDF and OWL recommendations, rdfschema recommendation
- Most widely used since for RSS-feeds (but not known for that. . .)
- My first contact with it: 2003

Basic concepts

“Everything should be representable, so one needs a common model with great generality”

*“Two basic elements:
Assertions
Quotations (statements about assertions)”*

`[http://www.w3.org/DesignIssues/Semantic.html]`

RDF model

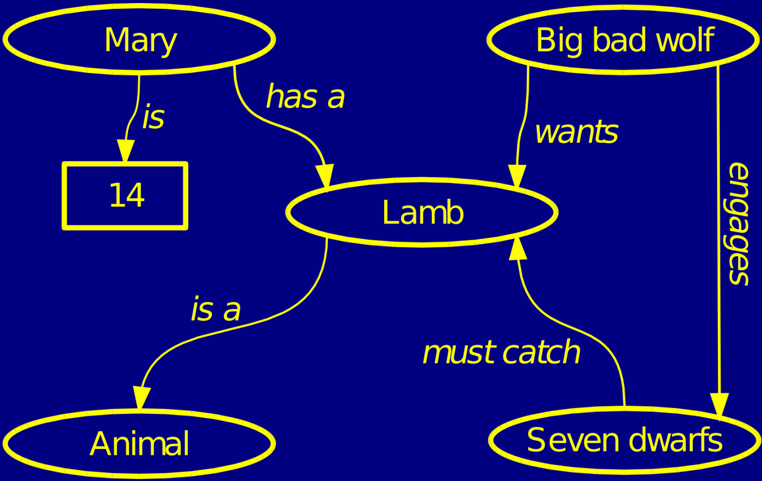
This leads to a very simplistic model, to RDF:

- Information is represented as a *triple*, as a statement
- Every triple consists of three elements:
  - **Subject**
  - **Predicate**
  - **Object**
- Subjects and predicates are given as URIs
- Objects can either be other URIs or literal data
- RDF data is represented by directed graphs

Example. . .

“Mary has a lamb.”
“Mary is 14.”
“Big bad wolf wants this lamb.”
“Big bad wolf engages the seven dwarfs.”
“The seven dwarfs must catch the lamb.”
. . . and of course. . .
“the lamb is an animal!”

Example - graph



Serialisation

And for the machines...

- **Notation 3 (N3)** - first form, quite complex
- **N-Triples** - subset of N3, possible recommendation
- **Turtle** - extension of N3
- **XML**

Example - turtle serialisation 1

```
➥ <http://example.org/Mary> <http://example.org/has>
<http://example.org/Lamb> .
➥ <http://example.org/Mary> <http://example.org/is> ‘‘14’’ .
➥ <http://example.org/BigBadWolf> <http://example.org/wants>
<http://example.org/Lamb> .
➥ <http://example.org/BigBadWolf> <http://example.org/engages>
<http://example.org/SevenDwarfs> .
➥ <http://example.org/SevenDwarfs>
<http://example.org/muststeal> <http://example.org/Lamb> .
➥ <http://example.org/Lamb> <http://example.org/isa>
<http://example.org/Animal> .
```

Note: lines are split up due to space. Each line starts with ➥

Example - turtle serialisation 2

Of course this isn't very handy, so here is a more cleaned up version:

- `@prefix ex:  <http://example.org> .`
- `ex:Mary ex:has ex:Lamb ;`
- `ex:is ‘‘14’’ .`
- `ex:BigBadWolf ex:wants ex:Lamb ;`
- `ex:engages ex:SevenDwarfs .`
- `ex:SevenDwarfs ex:muststeal ex:Lamb .`
- `ex:Lamb ex:isa ex:Animal`

Note: Turtle provides also other shortcuts not shown here

Example - XML serialisation

Example: XML-Serialisation. . .

RDF - additional features

- Typed literals: literals are identified by an URI. Example: XMLSchema-datatypes
- Localisation of literals (language specific literals)
- Empty nodes for complex relationships
- Containers, lists, collections

rdfschema - simple ontologies

rdfschema can be used to describe simple **ontologies**.

Ontology???

Taxonomy???

rdfschema - simple ontologies

- Basic mechanisms for structuring RDF data
- often already sufficient for simple ontologies
- Example: Dublin Core

Features of rdfschema

The most important constructs given by rdfschema are:

- `rdfs:Class`, `rdfs:subClassOf`
- `rdfs:Property`, `rdfs:subPropertyOf`, `rdfs:range`, `rdfs:domain`
- `rdfs:type`
- `rdfs:Container` (used for lists, sequences etc.)

rdfschema - examples

```
ex:Person rdfs:subClassOf <http://genome.org/human>

ex:Mary rdfs:type ex:Person

ex:is rdfs:subPropertyOf
<http://older.net/ageproperty>

ex:is rdfs:range ex:Person
```

OWL - introduction

OWL is the **W**eb **O**ntology **L**anguage

OWL - introduction

wait...
wouldn't that be **WOL** and not **OWL**???

OWL - introduction

Yes, but according to W3C:

- It's clear how to pronounce OWL. . .
- This acronym is great for making logos. . .
- OWLs are associated with wisdom. . .
- This makes up a great backstory. . .

[http://lists.w3.org/Archives/Public/www-webont-wg/2001Dec/0169.html]

A theory is also, that this comes from Winnie the Pooh, where the owl wasn't able to write her name correctly. . . ;-)

OWL - variants

OWL Full

- Contains **OWL DL** and **OWL Lite**
- Very expressive
- Not decideable (that causes headaches to reasoners. . .)
- Not fully supported by software

OWL DL

- Contains **OWL Lite**
- Decideable
- Nearly fully software supported

OWL Lite

- Decideable
- Fully software supported
- Less expressive

OWL Lite - constructs

OWL Lite Provides lots of constructs like

- (In-)Equality
- Property restrictions
- Cardinality restrictions
- Datatypes
- *. . . and others. . .*

Of course, all of this can be combined with rdfschema!

OWL DL/Full - constructs

OWL DL and OWL Full introduce even more constructs:

- oneOf
- disjointWith
- equivalentClass
- *. . . and others. . .*

Again - combineable with OWL Lite and rdfschema!

Full example

Very simple example taken from wikipedia. . .

And now...?

"It's very simple -
you read the protocol
and write the code."
(Bill Joy)

APIs

For programmers...
...Almost everything you need is already there...

- APIs
- Triplestores (RDF-Databases)
- Query-engines
- Reasoners
- *...and more!*

Check the linklist on the website of this speech for some hints...

Linking open data

- The "web of data", "giant global graph"...
- W3C initiative
- Consists of interlinked, openly accessible data
- 4.7 billion RDF triples, 142 milions of links
- **That's a huge amount of machine-readable knowledge already**

Browse and query it!

Tabulator demo

Infering and Reasoning

Besides of these uses, also infering and reasoning are interesting applications of RDF.

You can understand those as making automatic assumptions about triples. As a (very simple) example, look at this:

When we know that all human beeings are born. . .
. . . and we know that Tim Berners Lee is a human beeing. . .
. . . we can automatically infer that Tim Berners Lee has been born!

Reasoning and infering go into artificial intelligence. I won't go any further here too - but it's a very interesting field and you can - again find a lot of information and tools on the web!

Questions

Are there any…

…Questions?!?

…Answers?!?

Thanks!

Thanks a lot for your interest!

Check out

`http://impressionet.ch/semwebspeech3`

to find all the information (soon)!