The semantic Web: prospects & challenges

Michael Wilson
Manager, W3C Office in UK & Ireland, CCLRC Rutherford Appleton Laboratory
CCLRC Rutherford Appleton Laboratory

- Synchrotron: X-ray, light – life sciences
- Satellite ground station – Earth observation
- Particle accelerator Neutrons – materials science
- Laser beam: Nuclear fusion
- CERN Tier 1 - Accelerator data Particle physics
W3C Web Architecture
## Comparing Web and semantic Web development

<table>
<thead>
<tr>
<th>Web</th>
<th>Semantic Web</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 1989 proposal</td>
<td>• 1996 proposal</td>
</tr>
<tr>
<td>• +2 yrs portable browser</td>
<td>• +3 yrs RDF agreed</td>
</tr>
<tr>
<td>• +5 yrs commercial browser, index, crawler</td>
<td>• +8 yrs browser, OWL, RDF Schema</td>
</tr>
<tr>
<td>• +6 yrs 73,500 servers</td>
<td>• +9 yrs demonstration applications</td>
</tr>
<tr>
<td>• +10 yrs 3% US GDP from web</td>
<td>• +10 yrs commercial DB support</td>
</tr>
</tbody>
</table>

**Revolution**

**Evolution**
The semantic Web

• 1996 – metadata to describe the data on the Web.
• 1998 – query languages, inference rules & proof validation
• 2001 – bring structure to the meaningful content of web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users.
• 2006 – The “Semantic Web” is an infrastructure extending the current Web for the interchange and the integration of data on the Web.
Semantic Web Architecture & Technologies – 2001
Semantic Web Architecture & Technologies – 2006

• To make data machine processable, we need:
  – unambiguous names for resources (that may also bind data to real world objects): URI
  – a common data model to access, connect, describe the resources: RDF
  – access to that data, join the web: SPARQL
  – common access to XML and RDF: GRDDL
  – define common vocabularies, ontologies: RDFS, OWL, SKOS
  – Hyperlinks and the semantic Web – RDF/A
  – Reasoning – Rule Interchange Format
GRDDL

• Gleaning Resource Descriptions from Dialects of Languages (GRDDL)
• mechanism to relate other XML syntaxes (especially XHTML dialects) to the RDF abstract syntax via transformations
• Example – HTML Snippet:
  <html>
  <head><title>Example of an HTML document</title>
  </head>
  <body>

Translates to RDF assertion:

<document.html> dc:title "Example of an HTML document".

• Further work in W3C to take GRDDL to a recommendation
SPARQL – RDF Query Language

SELECT ?cat ?val # note: not ?x!
WHERE { ?x rdf:value ?val. ?x category ?cat }

Returns: ["Total Members",100],["Total Members",200],…,["Full Members",10],…]

• Waiting on XQuery to go to recommendation before SPARQL can.
SPARQL – Application
RDF-A

• RDF/A is a set of attributes used to embed RDF in XHTML.
• Instead of:

  If you want to contact me at work, you can either
  <a href="mailto:jo.lambda@example.org">email me</a>,
  or call +1 777 888 9999.

Use:

  If you want to contact me at work, you can either
  <a rel="foaf:mbox" href="mailto:jo.lambda@example.org">email me</a>,
  or call <span property="foaf:phone">+1 777 888 9999</span>.

To produce RDF triples:
  foaf:mbox = "mailto:jo.lambda@example.org"
  foaf:phone = "+1 777 888 9999"
Layering of the semantic Web

• Each layer should meet a new market need
• Each layer should build on previous ones
• Each layer should provide ROI by itself
Commercial RDF support

- Oracle v10.2
- Native RDF store
- Converters / loaders for existing RDF data
- RDF Query – An RDF_MATCH function which can be used in SQL to find graph patterns in RDF (similar to SPARQL)
- Over 100 Partners Thoroughly Test and Support Oracle Database 10g Release 2
Small Applications

Small apps in RDF provide ROI:

• RSS/Podcast
• Dublin Core
• XFN: bloggers
• Calendar
• Geo URL
• FOAF
• Thesaurus
• XML & RDF

DMOZ is an open directory project represents the directory an its contents in RDF.
Applications: Data Integration

- No "one schema" that can be used for describing everything
- No "right way" for describing / organizing anything
- Importance of "Partial Understanding"
- Things change, plan up front for it
- The value in "as needed" data integration
- Big wins come from many little ones
- The power of links – network effect
- Open-world, open solutions are cost effective
Application: data integration

- Semantic integration of different data sources
- RDF/RDFS (possibly with OWL and/or SKOS) based vocabularies as an “interlingua” among system components
- Many different projects and R&D on this: Boeing, MITRE Corp., Elsevier, EU Projects like Sculpteur and Artiste, national projects like MuseoSuomi...
Applications: Portals

- Vodafone's Live Mobile Portal
- search application (e.g. ringtone, game, picture) using RDF
  - page views per download decreased 50%
  - ringtone up 20% in 2 months
- Sun’s SwordFish: public queries for support, handbooks, etc, go through an internal RDF engine for White Paper Collections and System Handbook collections
- Nokia has a somewhat similar support portal
Common Themes to successful SW applications

- No "one schema" that can be used for describing everything
- No "right way" for describing / organizing anything
- Importance of "Partial Understanding"
- Things change, plan up front for it
- The value in "as needed" data integration
- Big wins come from many little ones
- The power of links – network effect
- Open-world, open solutions are cost effective
Summary

• Evolution not the revolution of the Web
• Small parts added to the architecture
• Development and adoption are progressing and continuing

But …
Research Challenges …
Six research problems for the SW

- Understanding ontological modelling
- Logical basis for inference
- Translating between ontologies
- Reasoning about agent’s intentions
- Sociology of agents
- Governance of agents
Understanding ontological modelling

• Guizzardi addressed some of these issues yesterday.
• OWL is an epistemological language for ontologies
• What is the development method for ontologies?
• What are the design languages?
• What are the design trade-offs in ontology development?
• What is the ontological status of objects and relations?
Logical basis for inference

• Are Description Logics insufficient for rule based reasoning?
• Are Horn clause logics required instead?
• The constraints on Horn clause logic and description logics are different.
• What are the consequences of using the two together?
Translating between ontologies

• Today
  – transform to common syntax
  – Take union of concepts & axioms
  – Define bridging axioms through an inference engine.

• Eventually
  – Metaphors are assertions of conceptual identity
  – Translation using metaphoric relations
Reasoning about agent’s intentions

- Certified Assertion based confidence
- Transfer/Generalisation of assertions?
- Generalisation as trust – common intention
- How do we transfer/generalise intentions?
Sociology of agents

• Agents operate in networks
• Agents cluster and herd
• What properties of the group behaviour of agents are important?
Governance of agents

- Liability and Governance
- Extending
  - the development method
  - the run time system
  - service quality
Conclusion

• The larger vision of the SW is still valid
• It will take a long time to evolve from the web of today into a full SW
• Much research in more abstract areas is required before we will get there
• The opportunities are there for you to take