

Current Status and Future Promise of the **Semantic Web**

James Hendler
University of Maryland
College Park, MD

Ora Lassila
Nokia Research Center
Cambridge, MA

November 2006



Before We Get Started...

- **Two views of the Semantic Web:**
 - implementing SEMANTIC applications using web technologies
 - using semantic technologies to support new WEB applications
- **Exploring the relationship (and *tension*) between the two over time**
- **This talk is**
 - a retrospective, a status report
 - an “interpretation”
 - some thoughts on the future...



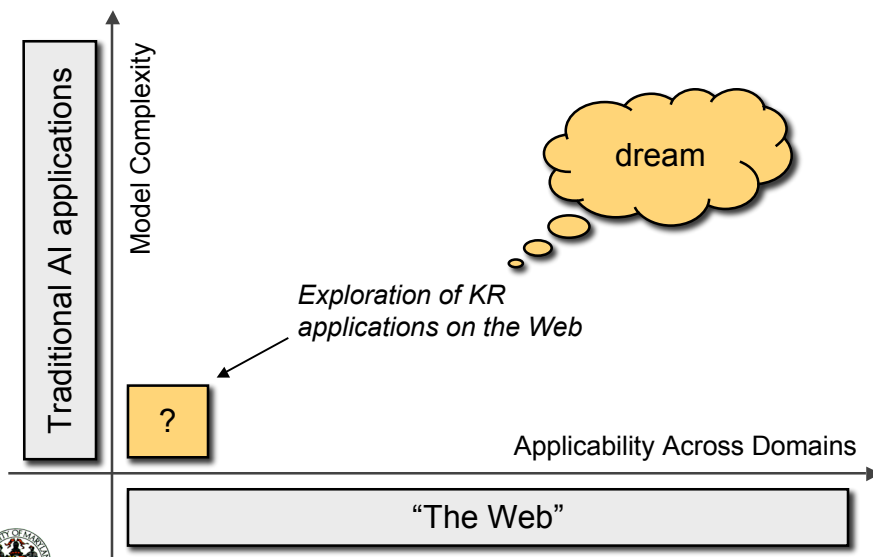
1990's: "Pre-history"

- **Rebirth of Artificial Intelligence (end of "AI Winter")**
 - "big" AI applications
 - Deep Blue, Mars Rover, Deep Space 1, ...
 - embedded vs. stand-alone
 - Web AI
 - IR, statistical NLP, machine learning
 - lots of data!
- **Emergence of the Web**
 - new ways of doing things
 - new business models (even new *social* models)
 - new technology
 - "dot-com" boom
- **Early forays into "meta-content"**



NOKIA
Connecting People

1990's: "Pre-history"



NOKIA
Connecting People

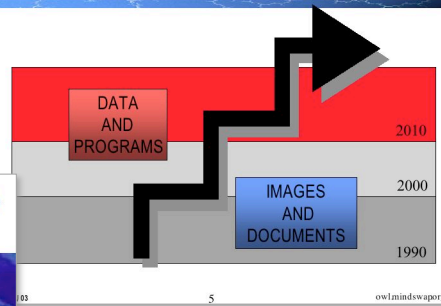
2000-2001: What Did We Believe?

- Jim: Semantic Web and the advent of pervasive computing (March 2000)

TetherlessDAML
The Mobile Semantic Web

<http://www.daml.org>

What about the rest of the Web?



- Jim: Roadmap from the "old" Web to the Semantic Web (October 2001)

NOKIA
Connecting People

2000-2001: What Did We Believe?

- Ora: Semantic Web and the advent of pervasive computing (June 1999)

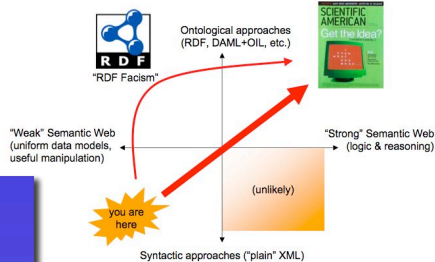
Why Your Microwave Oven Needs (to Surf) the WWW

Ora Lassila

Research Manager
Agent Technology Group, Nokia Research Center / Boston
&
Member of the Advisory Board
World Wide Web Consortium (MIT/LCS)

June 1999

Semantic Web: Characterizations



- Ora: Roadmap from the "old" Web to the Semantic Web (October 2001)

NOKIA
Connecting People

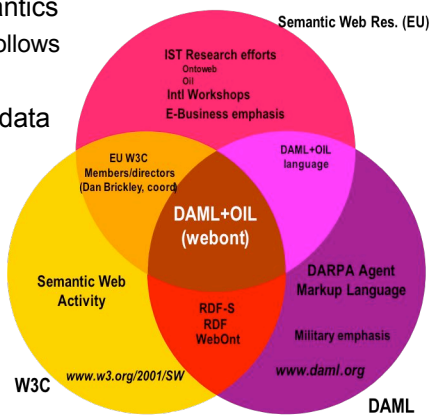
2000-2001: “Early Years”

- “Dot-Com” optimism still prevails:
easy to explore new directions

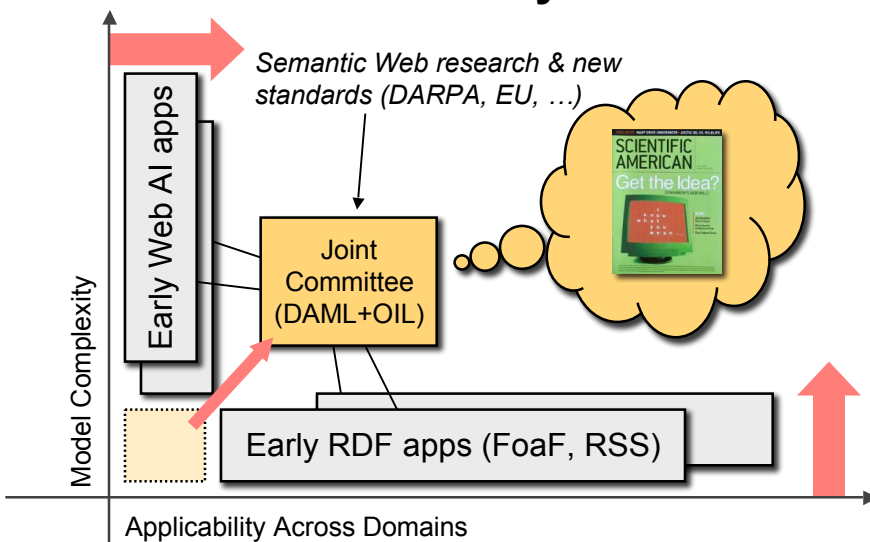
- Government meddles with semantics
 - DARPA’s DAML program; EU follows
 - DAML+OIL
- Web community discovers metadata
 - W3C Metadata Activity
 - RDF

“3-pronged” attack:

- DARPA
- EU IST
- W3C



2000-2001: “Early Years”



Original Outline (July 2000)

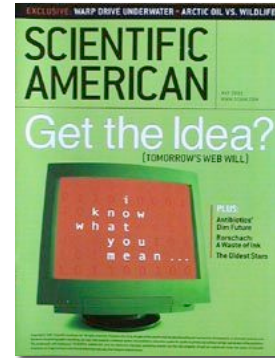
Scientific American Article notes

[Joint starting place:]

- I. Semantic Web Vision (TBL)
- II. What are the enablers? (in sequence)
 - Screen Scraping (Ora and TBL)
 - Data on Web (Ora and TBL)
 - Zip code link between Data Bases (TBL)
 - Ontology Independence (JAH)
 - Effect of Scale (TBL)

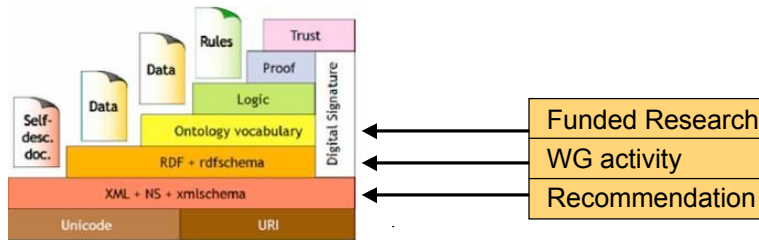
“Then, a miracle occurs”

- III. What can you do with it? (not necessarily in sequence)
 - Self-describing documents (JAH)
 - Logic to encode... (TBL)
 - Services and Advertising (Ora)
 - Devices (Ora)
 - Digital Signatures, Authentication, and Trust (TBL)



NOKIA
Connecting People

2001



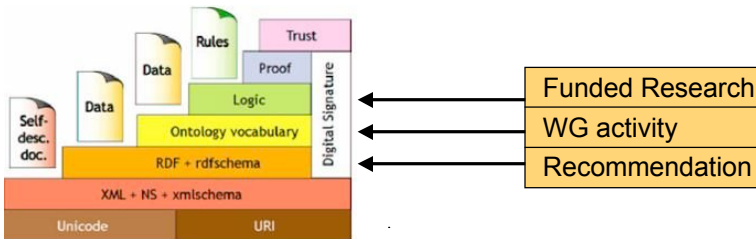
Semantic Web Today

- The Semantic Web of 2002 resembles the early days of the World Wide Web
 - ◆ Development funded primarily by Govt, but emerging corporate interest
 - ◆ A lot of excitement, but confusion as to business case
 - ◆ Open source tools and geeks in control
 - ◆ Standards starting to stabilize to point where they permit deployment
 - ◆ Developer tools, libraries, languages

- Research, experimentation, early demonstrations
- Reminiscent of the early days of the Web



2003



Semantic Web Today

- The Semantic Web of 2002 research World Wide Web
 - Development funded primarily by Gov
 - A lot of excitement, but confusion as well
 - Open source tools and "geeks in control"
 - Standards starting to stabilize to point of use
 - Developer tools, libraries, languages

10

"Our" Semantic Web

- Jan 1, 03: Crawler finds 5.8M+ DAML statements on 20,000+ web pages
 - Doesn't include many instance KBs tied to ontologies
 - Doesn't include many very large RDF S-based KBs that include some OWL
- Ontology library at <http://www.daml.org> has 195 ontologies (March 2003)
 - Open for anyone to create
 - Open for anyone to use
- OWL is being supported by large corporation labs
 - Web tool developers: IBM, HP, Sun, Intel, Fujitsu
 - Content providers: Daimler-Chrysler, Nokia, Motorola, EDS, Agfa
- OWL is starting to be used by thesaurus developers
 - C.I. National Cancer Institute: metathesaurus released in OWL, Lite
 - United Nations Standard Product Codes available in DAML
 - NASA thesaurus available in DAML
- Use of semantic markup for Web Services beginning to move beyond basic research
 - DAML-S cited as required reading for Web Services Composition WG

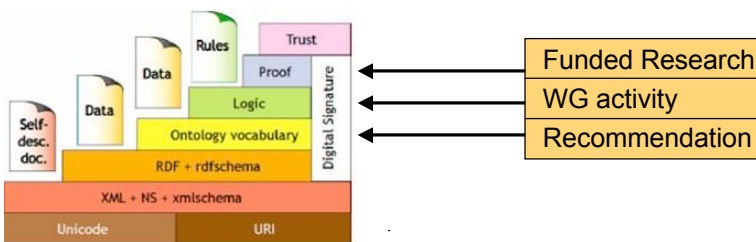
Sanken, 03

23



- Early government adoption
- Emerging corporate interest

2005



Semantic Web Today

- The Semantic Web of 2002 research World Wide Web
 - Development funded primarily by Gov
 - A lot of excitement, but confusion as well
 - Open source tools and "geeks in control"
 - Standards starting to stabilize to point of use
 - Developer tools, libraries, languages

10

"Our" Semantic Web

- Jan 1, 03: Crawler finds 5.8M+ DAML statements on 20,000+ web pages
 - Doesn't include many instance KBs tied to ontologies
 - Doesn't include many very large RDF S-based KBs that include some OWL
- Ontology library at <http://www.daml.org> has 195 ontologies (March 2003)
 - Open for anyone to create
 - Open for anyone to use
- OWL is being supported by large corporation labs
 - Web tool developers: IBM, HP, Sun, Intel, Fujitsu
 - Content providers: Daimler-Chrysler, Nokia, Motorola, EDS, Agfa
- OWL is starting to be used by thesaurus developers
 - C.I. National Cancer Institute: metathesaurus released in OWL, Lite
 - United Nations Standard Product Codes available in DAML
 - NASA thesaurus available in DAML
- Use of semantic markup for Web Services beginning to move beyond basic research
 - DAML-S cited as required reading for Web Services Composition WG

Sanken, 03

23

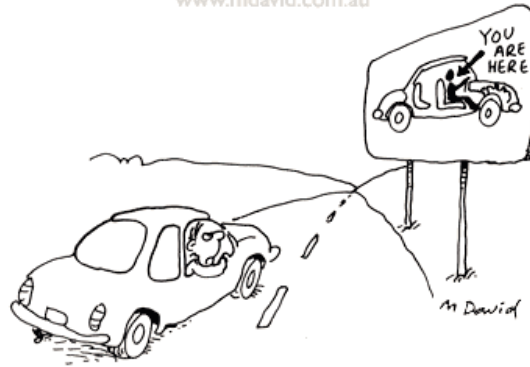
Now

- Companies getting into the act
 - IBM SNObase ontology management system
 - Oracle to support RDF in Oracle 10.2
 - Open source (Kowari) scalable triple store
 - Supports RDFS
 - OWL support coming
 - HP Labs open-source Jena API
- Many open source tools becoming available for experimentation/academic use
 - Kowari, RDFLib, 3Store...
 - Jena, Sesame...
 - Protégé, SWOOP, Onto(xxx)...
- Building corporate demonstrators becoming cheap and easy

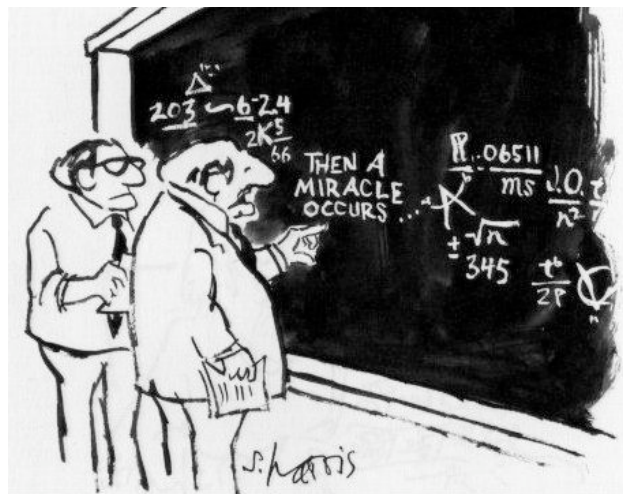


2006: You Are Here!

www.mdavid.com.au



Then a Miracle Occurs...



"I think you should be more explicit here in step two."

from *What's so Funny about Science?* by Sidney Harris (1977)

Significant Corporate Activity

- **Semantic (Web) technology companies starting & growing**
 - Siderean, SandPiper, SiberLogic, Ontology Works, Intellidimension, Intellisophic, TopQuadrant, Data Grid, ...
- **Bigger players buying in**
 - Adobe, Cisco, HP, IBM, Nokia, Oracle, Sun, Vodaphone... announcements/use in 2005-2006
 - integrator and contractor uptake: Northrop Grumman buys Tucana, Lockheed-Martin uses SiberLogic in FCS, WebMethods buys Cerebra, ...
 - tools being announced: AllegroGraph, TopBraid, ...
- **Government projects in and across agencies**
 - US, EU, Japan, Korea, China, ...
- **Life sciences/pharma an increasingly important market**
 - Health Care and Life Sciences Interest Group at W3C
- **Many open source tools available**
 - Kowari, RDFLib, Jena, Sesame, Protégé, SWOOP, Onto(xxx), Wilbur, ...



NOKIA
Connecting People

Significant Corporate Activity

semantic web company - Google News

http://news.google.com/news?hl=en&ned=us&ie=UTF-8&q=semantic+web+company

Results 1 - 10 of about 74 for semantic web company. (0.09 seconds)
Sorted by relevance Sort by date

50+ Semantic Web press releases within one month!

Caboodle Networks Launches Open Source Semantic Web-Based Search ...
LinuxElectronics, TX - Mar 4, 2006
... Networks has launched an open source search engine based upon semantic web standards ... to discovery by traditional text-retrieval search engines the company claims ...

Access-eGov: Accessible eGovernment embracing the Semantic Web
PR Leap (press release), CA - Mar 2, 2006
... to e-Government Services Employing Semantic Technologies) is ... get a building permit, register a new company, (web) accessibility expertise ...

Hooking up with RDF
GON.com, DC - Mar 3, 2006
... the Resource Description Framework in his 2003 book The Semantic Web, but for ... available with the Oracle Spatial the company's software ...

Mandriva to help build "Social Semantic Desktop"
Desktop Linux - Feb 28, 2006
... Social Semantic Desktop represents ... The company will also take over the submission of NEPOMUK ... standard the World Wide Web Consortium, as ...

News For Your Web Site
dBusinessNews San Jose (press release), CA - Feb 23, 2006
... National Institutes of Health, Raytheon, The Walt Disney Company, and General ... as other respected innovators such as Ericsson, the Semantic Web activity lead ...

Hip hop offers data integration lessons
InfoWorld, Netherlands - Feb 22, 2006
... use information that is available on the Web or in company databases and some ... for mashups and other applications would be to use Semantic Web technology such ...

Parasoft Releases Latest Version Of Automated Test Software For ...
SYS-CON Media, NJ - Mar 2, 2006
According to the company, the new release delivers extensive ... they need to validate their web services for ... fall into four categories: BPEL semantic tests, WSDL ...

Experiments in data integration
Webwereld, Netherlands - Feb 23, 2006
... The company evolved from a group of University of Washington scientists and has rapidly ... XDA is due in early 2006 that will incorporate Semantic Web technologies ...

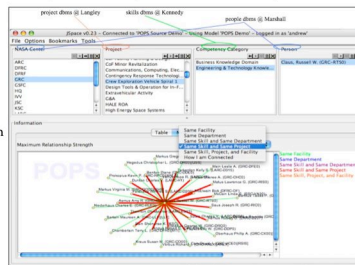
Finally, there is an Alternative to the Internet
PR.com (press release), NY - Feb 23, 2006

Growing Government Activity (US&EU)

- Agencies moving beyond the "talk" phase
 - primarily prototyping, but first acquisitions starting
- Example:
 - NASA is developing an enterprise data strategy around using existing data via Semantic Web integration

Lots of activities across NASA

- Science, Engineering, and Mission all have SWT production or development efforts in place
- Now focus is on re-using the data systems we already have in place
- Agency wide integration planning is underway for building a federation of models into an integrated information service across all disciplines



(A. Schain, 3/06)



There's a Lot Out There!

2,120,000 hits on "rdf filetype:rdf"

13,600 hits on "ontology filetype:owl"

Paid ads



(March, 2006)

Swoogle Semantic Web Search Engine

http://swoogle.umbc.edu/index.php?topoo...

Prof. James A. Hendler World Wide Web Con... RDF Interest Group L... Planet RDF Google Scholar

Swoogle

Swoogle's Statistics of the Semantic Web

MAIN MENU

- swoogle search
- swoogle statistics
- swoogle feedback
- submit url

DOCUMENTATION

- publications
- manual
- news

LINKS

- project home
- swoogle 2005

LOGIN FORM

Username

Password

Remember me

Login

Lost Password? No account yet? Register

SQL SITESTATS SUMMARY

Site Stats Summary

	Hits	Visitors
Total	549320	44484
Today	1417	296
Week	14903	1945

news | faq | feedback | web-service | submit-url | swarchive | swoogle-2005

Copyright © 2004-2006, Ebiquity Group at UMBC
This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License.

<http://swoogle.umbc.edu>

More OWL Use

- The OWL namespace has been declared by 113,000 SWDs (8%) and actually used by 108,000 (7%)
- The RDFS namespace enjoys more use, being declared by 677,000 (47%) and used by 538,000 (37%) SWDs
- owl:Class is the most used term from the OWL namespace with about 1,800,000 instantiations in 68,000 SWDs
- significant use of two OWL equality assertions: owl:sameAs (280,000 assertions in 17,000 SWDs) and owl:equivalentClass (70,000 assertions in 4,300 SWDs) – their common use may be an indication of increased ontology alignment

(from Ebiquity blog, Sept 1, 2006)



Semantic WEB

My Playz Bank

Map

Options

Data harvesting & visualization

A little **Semantics** goes a long way



Locations

- Amherst
- Bedford MA
- Boston Logan
- Burlington N
- Calgary AB

Options

Options

EXIT

Rich metadata

February 06, 2006

A Clockwork Orange

A few nights ago, my friend Yvonne and I went to one of the local theaters where they're currently having a classic film festival and saw *A Clockwork Orange*. For never seen it on the big screen before, but after viewing it, I realized just how much of an amazing of a movie it is, and how much of a master genre *Stanley Kubrick* was.

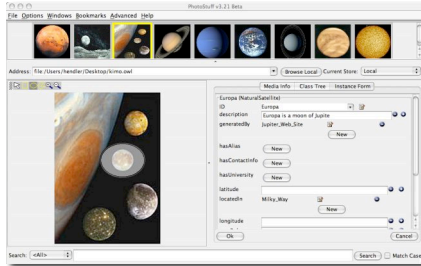
Even of the movie is nearly 35 years old, the idea of its premise still gives the chills. Set in a future Britain where civil society is on the brink of total collapse, young Alex and his gang of thugs rightly name the city—presumably London—having people up with no names, hoping to

This page mentions the movie *A Clockwork Orange* (1971). The average rating for this film is 3 stars (out of 4). Your recommended rating from the FirstTrust system is 0.5 stars.

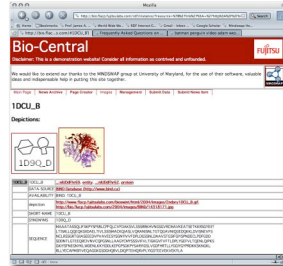
Web-based social networks



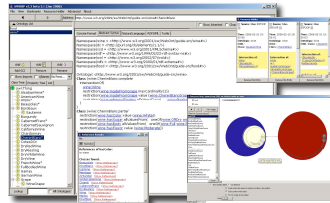
SEMANTIC Web



Digital asset management



Scientific portals



Tools for developers

A little Web
goes a long way



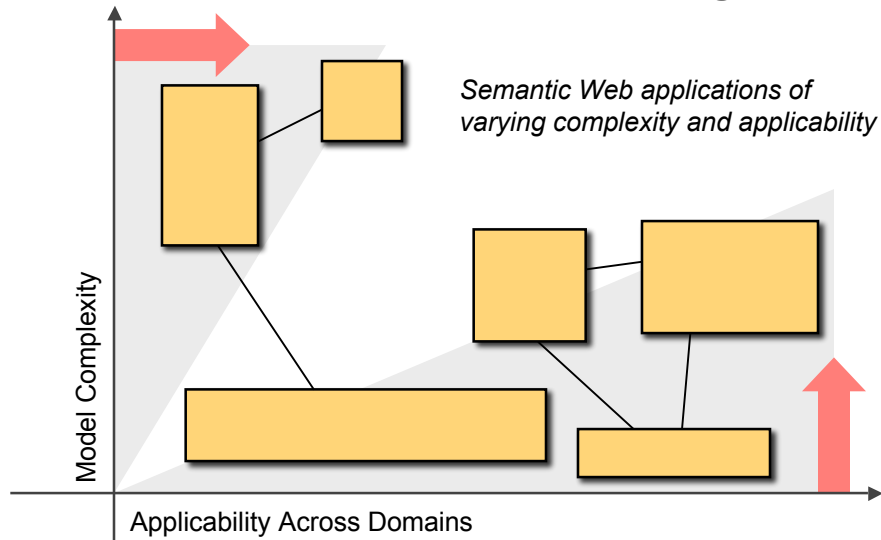
Enterprise Information Integration

- **Deployment of semantic technologies is easier in a “controlled” environment**
 - such as a corporate intranet
- **Key benefits from Semantic Web Technology:**
 - reuse of installed clients and servers
 - careful design of SW languages for Web compatibility
 - leave data in place, integrate through an RDF store
 - analogous to 3-tiered Web application
 - heterogeneity supported by ontologies

"Corporate Semantic Web", Gartner "hot pick" for 2006



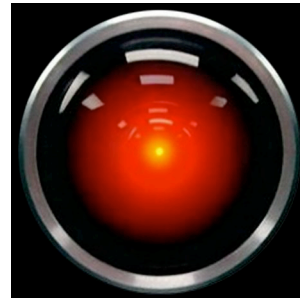
2006: The Gap Is Closing



NOKIA
Connecting People

SEMANTIC Web Lessons

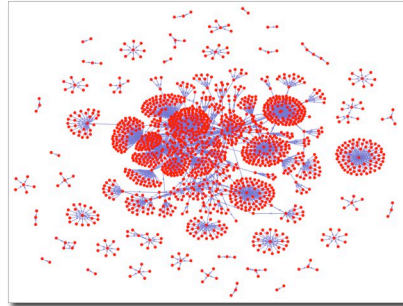
- **What we learned from AI...**
 - embedded AI succeeded, stand-alone did not
 - tools are hard to sell
 - reasoners are a *means*, not an end
 - knowledge engineering bottleneck
- **...applied in the Web context**
 - futureproofing
 - URIs are important
 - good standards evolve
 - languages (RDFS, OWL, RIF, ...)
 - content!



NOKIA
Connecting People

Semantic WEB Lessons

- **Web needed high value sites**
 - personal (homepages, pets)
 - public (hobbyists, govt)
- **As these linked up, new functionality emerged**
 - Yahoo, Alta Vista, ...
- **New business models followed...**
 - “give it away” (Netscape)
 - marketplace (Amazon)
 - advertising (Yahoo, Google)
- **Semantic Web?**
 - **SHARE; GIVE IT AWAY!**

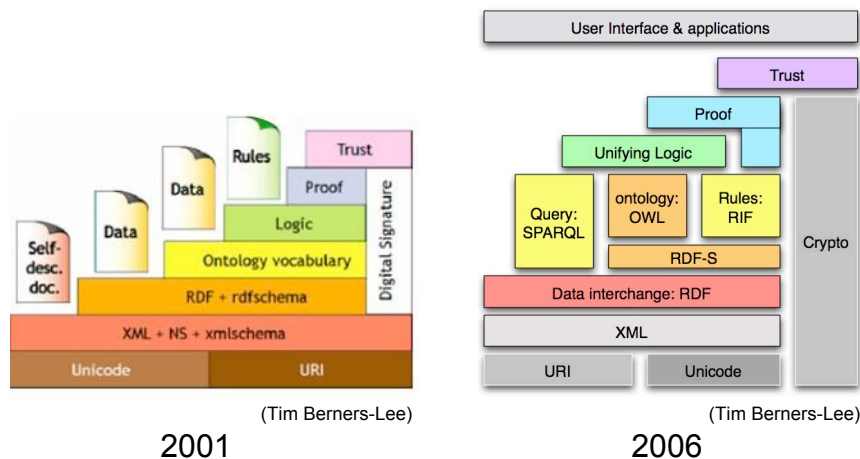


- **What do we need?**
 - Open Source Tools
 - Open Source Datasets
 - Open Source Harvesters



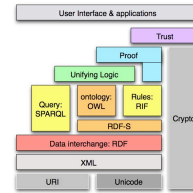
NOKIA
Connecting People

The “Layer Cake” is Evolving...



NOKIA
Connecting People

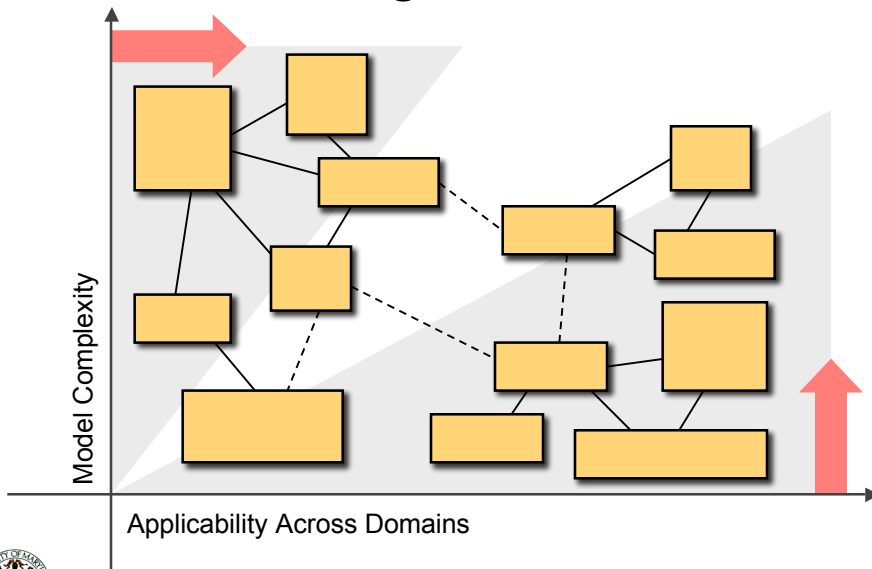
New Languages Underway



- **RIF: Rules Interchange Format**
 - representing rules on the Web
 - linking rule-based systems together
- **SPARQL: Query language for (distributed) triple stores**
 - the “SQL of the Semantic Web”
- **GRDDL/RDFa: Integration of HTML and Semantic Web**
 - “embedding” RDF-based annotation on traditional Web pages
- **OWL: New features, specialized subsets**
 - RDF++/OWL Mini – simplification, identity, scaling to large datasets
 - OWL 1.1 – additional expressivity for OWL constructs
- **And more...**
 - multimedia annotation, Web-page metadata annotation, Health Care and Life Sciences (LSID), privacy, etc.



Linking Is Power!

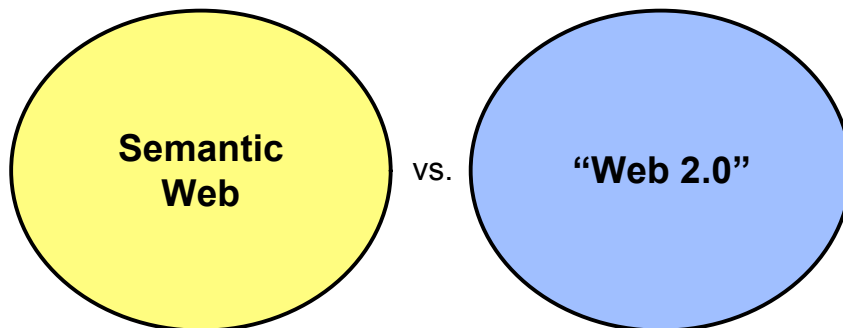


Semantic Web vs. “Web 2.0”

- **Data with formal semantics**
 - RDF, OWL
 - SPARQL, RIF
 - **Spontaneous information integration (finally!)**
 - **Semantic Web services, agents**
 - **Strong emphasis on open standards**
- **New social phenomena: blogs, wikis, tagging, folksonomies**
 - **New user interfaces**
 - AJAX (or: “Rich User Experience”)
 - **“New” kinds of data**
 - microformats, RSS
 - “mash-ups”
 - Web services
 - **Plays “fast & loose” with standards**

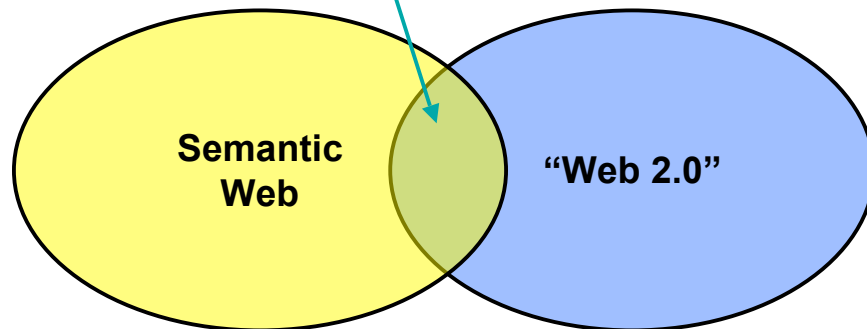
Semantic Web & “Web 2.0”

- What is their relationship?
- Will they stay separate? Does that even make sense?



Semantic Web & “Web 2.0”

- **NO!** Considerable synergies exist



NOKIA
Connecting People

Exploiting “Web 2.0”

- **Vast amounts of “semi-engineered” knowledge**
 - Flickr: tens of millions of keyword-tagged photos
 - microformatted Web documents
 - Wikipedia: thousands of carefully documented subjects (in a hierarchy, with disambiguation, ...)
- **Generate “persistent” URIs**
 - “Tank” <http://en.wikipedia.org/wiki/Tank> (armament)
 - “Tank” http://en.wikipedia.org/wiki/Tank%2C_Pakistan (small town in Pakistan)
- **Remember: Anything with a URI can be linked to the Semantic Web!**



NOKIA
Connecting People

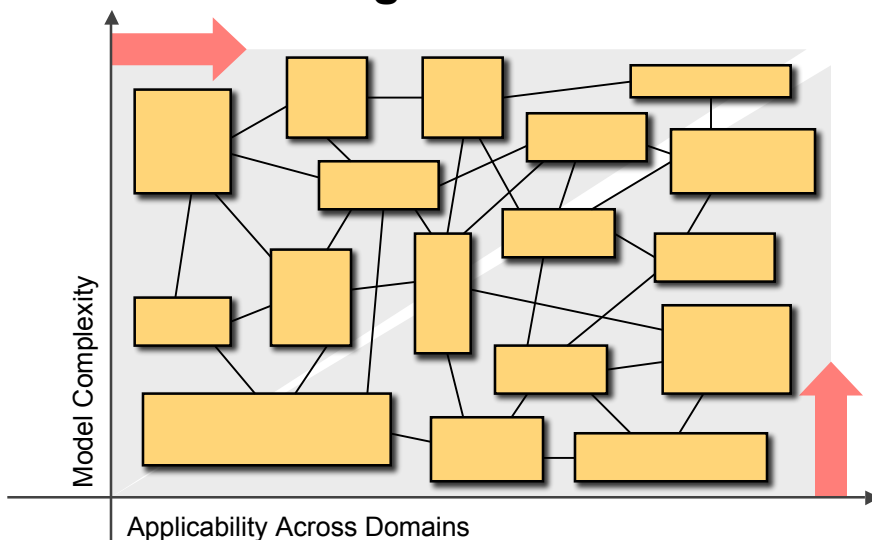
Linking of "Web 2.0" & Semantic Web

- Using informal Knowledge Engineering (KE) to bootstrap "formal" KE
- Extending formal KE from tags/wiki

The screenshot shows a browser window with two tabs. The left tab displays an OWL class named 'cat' with various properties and instances. The right tab shows the Wikipedia article for 'Cat', which includes a detailed description, a scientific classification table, and an image of a cat. The OWL class is linked to the Wikipedia article, demonstrating the integration of formal and informal knowledge.

Scientific classification	
Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia
Order:	Carnivora
Family:	Felidae
Genus:	Felis
Species:	<i>F. catus</i>

Looking Further Out



Where Are the Agents?

- **“Brave New Applications”**
 - operate autonomously in “unanticipated” situations
 - exhibit robustness in the face of
 - changing, inconsistent and unexpected data
 - variations in reliability, trust
 - capable of serendipitous behavior, opportunism
- **Move from the “tool use” of personal computing to systems that *work on our behalf***
- **(Semantic) Web services as “plumbing” for agents**
 - emerging as we speak...



NOKIA
Connecting People

Pervasive Computing & Semantic Web

- **Pervasive Computing is an interoperability nightmare!**
 - instead of sometimes connecting a handful of devices, dynamically connect/disconnect/reconnect possibly hundreds of devices
- **Today, high cost of ensuring interoperation**
 - any interaction has to be specifically designed/engineered
 - heavy emphasis on application-specific standardization
 - spontaneous interoperability is next to impossible
- **The vision is largely contingent on getting unanticipated “encounters” of devices to work**
 - how do you behave in a situation not covered by a standard?
 - not “future-proof”
- **Semantic Web is a good match**
 - (it is an “interoperability technology”)



NOKIA
Connecting People

Other Emerging Trends

- **Semantic Web Services**
 - crucial for linking “programs” into the mix
 - “plumbing” for agents...
- **Scaling Semantic Web stores to database sizes**
- **Information extraction and semantics (“Web 3.0”)**
 - can we “retrofit” semantics on the existing Web?
- **Semantic Web information creation**
 - can we make it so we don’t have to retrofit in the future?
 - tools that help embed the semantics as a document is created
 - better dynamic integration of structured data into the Semantic Web
 - “Semantic Desktop”



Summary

- **Most things we predicted have happened**
 - (or are happening at the moment...)
- **Some things happened faster than we anticipated**
 - triple store scaling
 - reasoner performance *actually matters*
 - ontologies are there (but very little linking)
- **Some things are yet to materialize (but we are hopeful)**
 - public information sources (as RDF, OWL, ...)
 - digital convergence, pervasive computing just emerging
 - little progress on agents

Now go out there and make some money off this...!



Any Questions?

“A Little Web Goes A Long Way”

Model Complexity



The Semantic Web

Applicability Across Domains

“A Little Semantics Goes A Long Way”

