

Semantic Web @ 5

Current Status and Future Promise of the Semantic Web

James Hendler
University of Maryland
College Park, MD

Ora Lassila
Nokia Research Center
Cambridge, MA

March 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”
 - a vision for the future

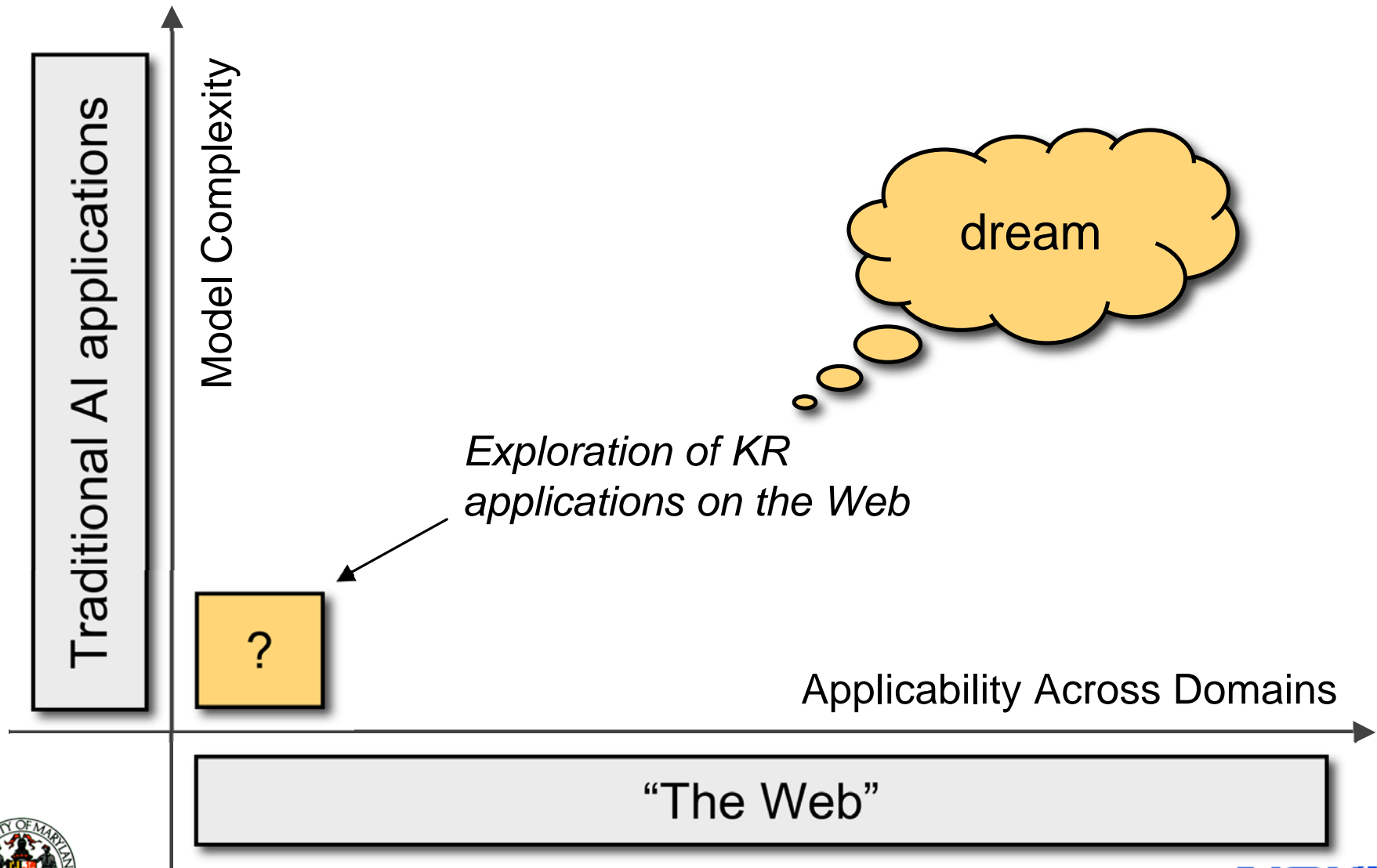


1990's: “Pre-history”

- **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”**



1990's: "Pre-history"



2000-2001: What Did We Believe?

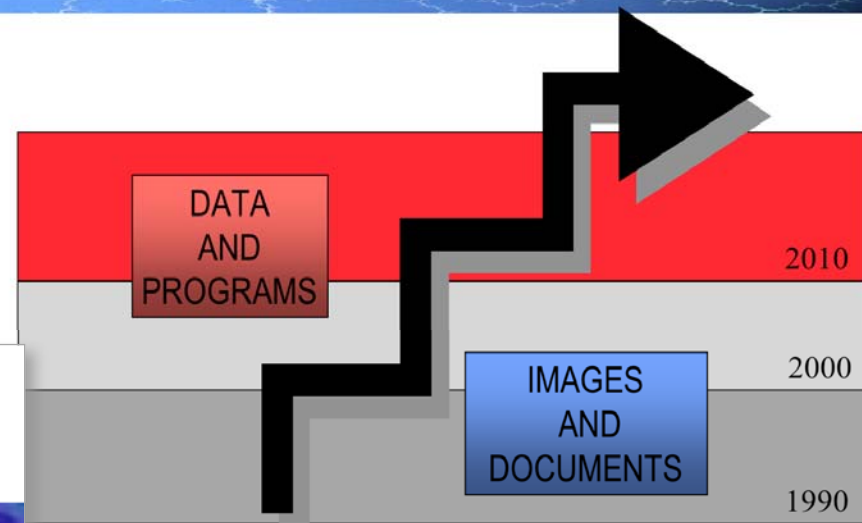
- **Jim: Semantic Web and the advent of pervasive computing (March 2000)**

 **TetherlessDAML** 
The Mobile Semantic Web



<http://www.daml.org>

What about the rest of the Web?



03

5

owl.mindswap.org

- **Jim: Roadmap from the “old” Web to the Semantic Web (October 2001)**

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

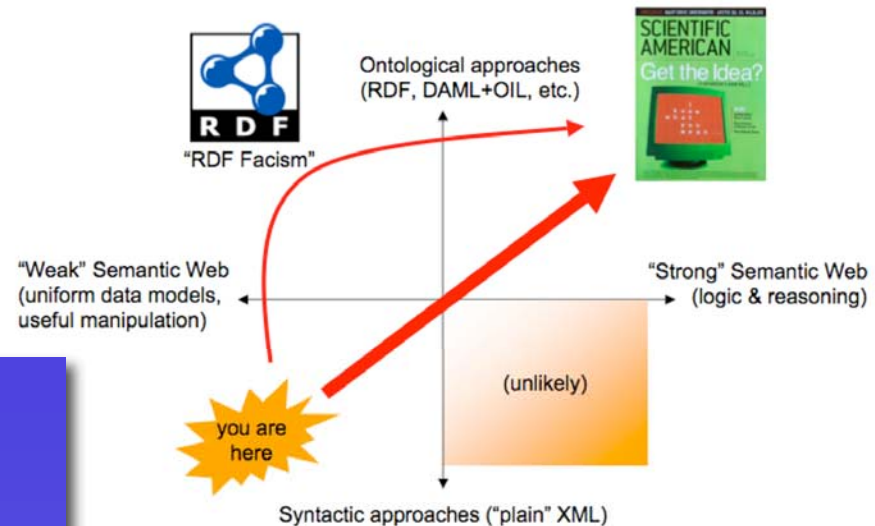
Ora Lassila

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



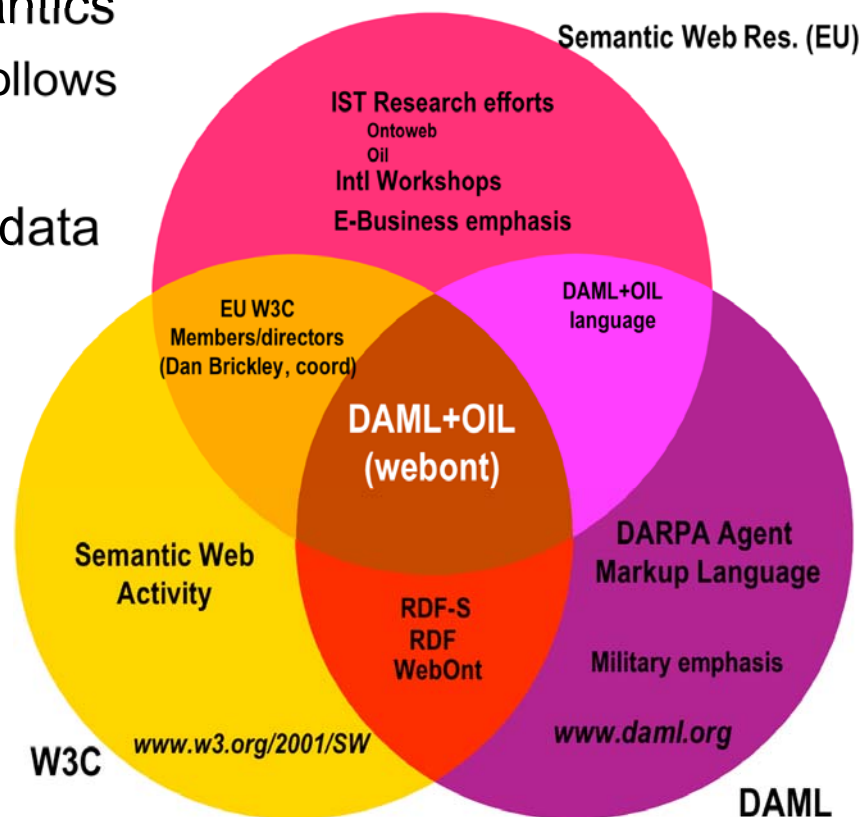
- **Ora: Roadmap from the “old” Web to the Semantic Web (October 2001)**

2000-2001: “Early Years”

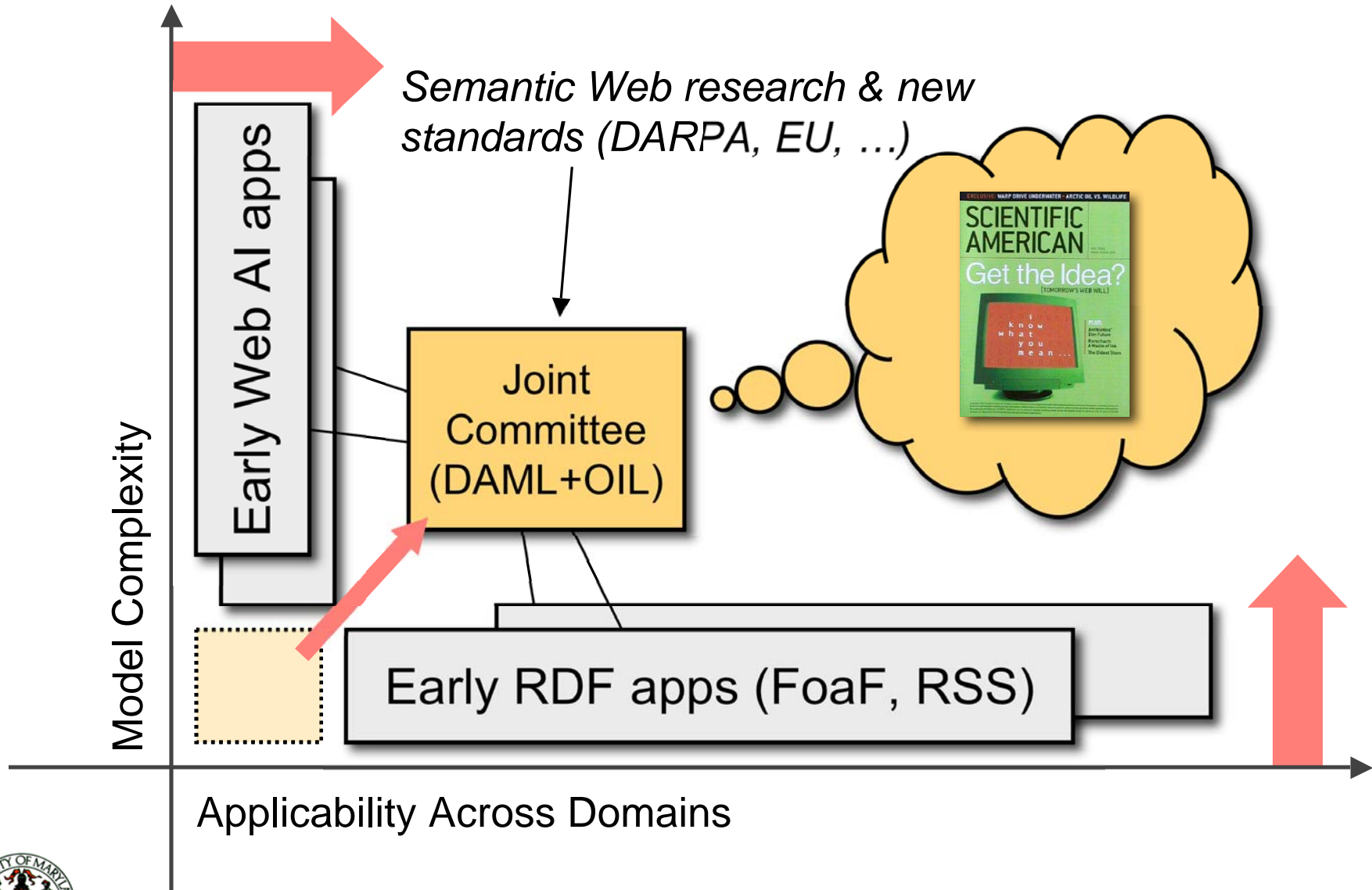
- “Dot-Com” optimism still prevails:
easy to explore new directions
- Government meddles with semantics
 - DARPA’s DAML program; EU follows
 - DAML+OIL
- Web community discovers metadata
 - W3C Metadata Activity
 - RDF

“3-pronged” attack:

- DARPA
- EU IST
- W3C



2000-2001: “Early Years”



Original Outline (July 2000)

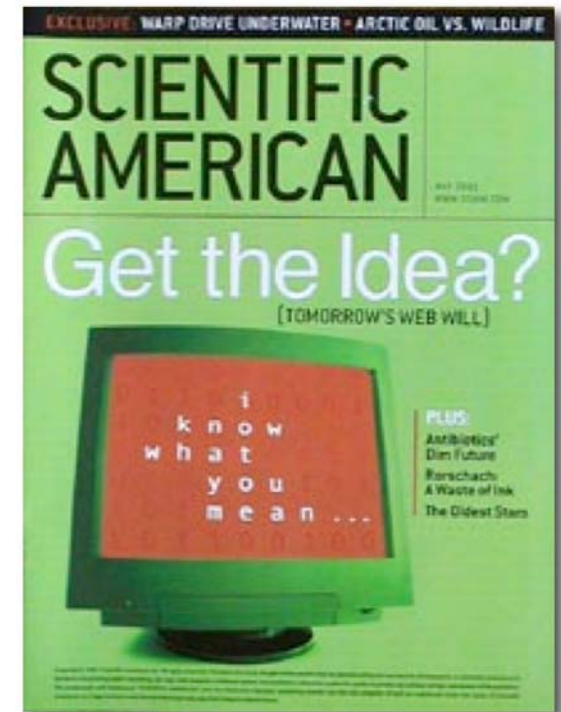
Scientific American Article notes

[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”

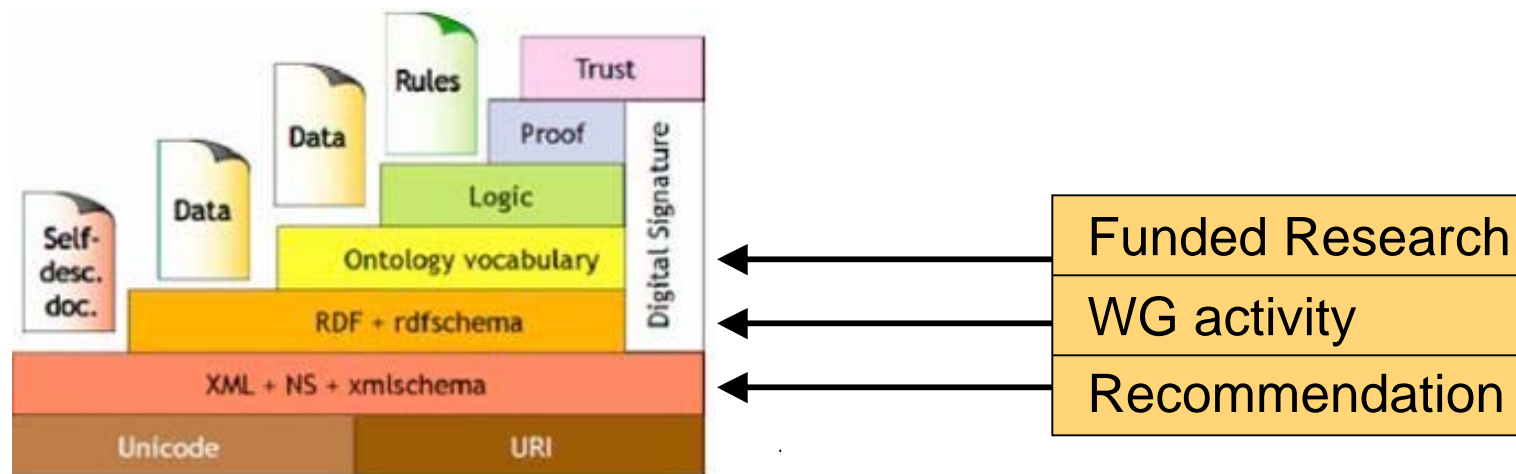
- III. What can you do with it? (not necessarily in sequence)
 - Self-describing documents (JAH)
 - Logic to encode... (TBL)
 - Services and Advertising (Ora)
 - Devices (Ora)
 - Digital Signatures, Authentication, and Trust (TBL)



(historical note: last draft to printers - March 9, 2001)



2001



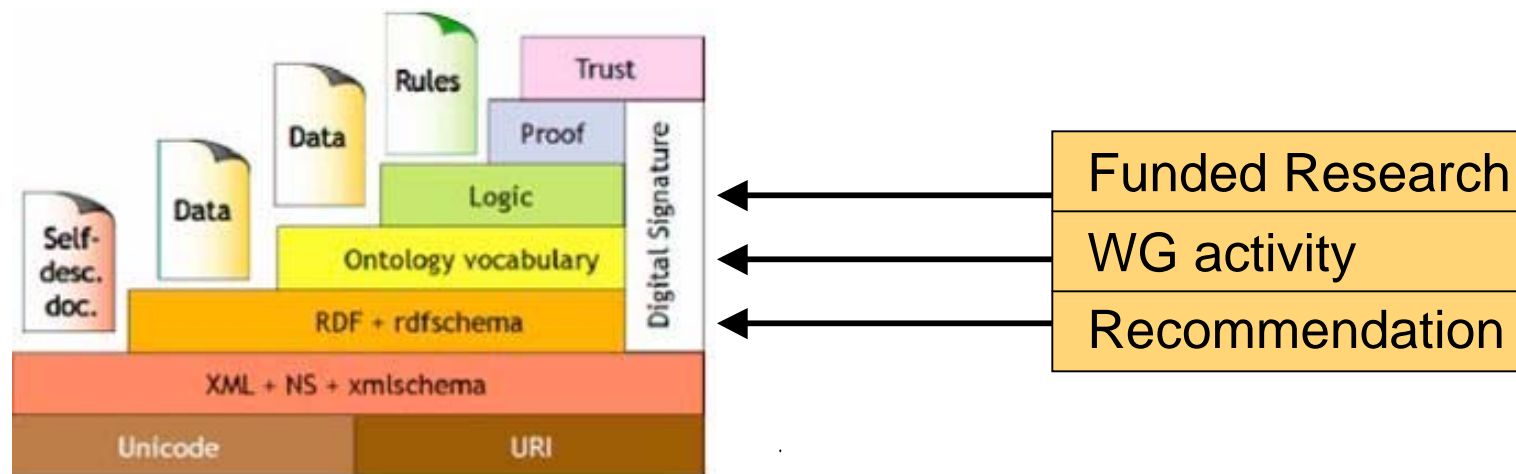
Semantic Web Today



- 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 resembles the World Wide Web
 - ◆ Development funded primarily by Government
 - ◆ 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 RDFS-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.f. 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

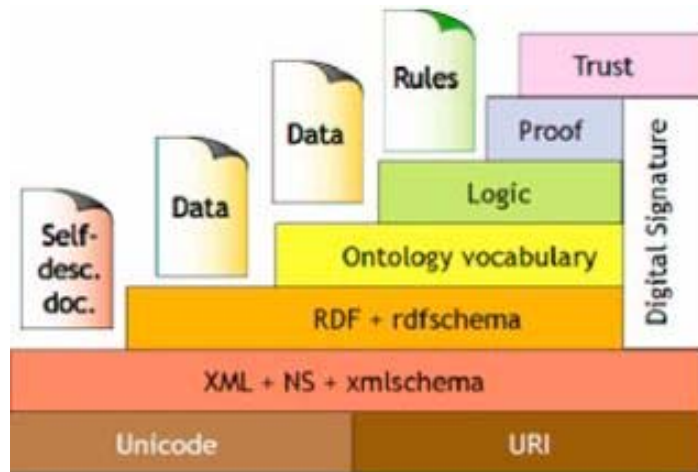
Sankar, 03

23

www.mindswap.org

- Early government adoption
- Emerging corporate interest

2005



Funded Research

WG activity

Recommendation

Semantic Web Today

- The Semantic Web of 2002 reser World Wide Web
 - Development funded primarily by Gov
 - A lot of excitement, but confusion as
 - Open source tools and "geeks in conf
 - Standards stag
 - Developer tools, libraries, languages

- **Commercial tools**
- **Lots of open source software**
- **Scalability**

"Our" Semantic Web

- Jan 1, 03: Crawler finds 5.8M+ DAML staten
 - Doesn't include many instance KBs tied to one
 - Doesn't include many very large RDFS-based
- Ontology library at <http://www.daml.org> has
 - Open for anyone to create
 - Open for anyone to use
- OWL is being supported by large corporati
 - Sun, Intel, Fujitsu, etc.
 - Content providers: Carmer, Chrysler, Nokia, M
- OWL is starting to be used by thesaurus de
 - National Cancer Institute metathesaurus re
 - News Standard Product Codes availa
 - NASA thesaurus available in DAML
 - For Web Services b
 - DAML-S cited as required reading for Web S

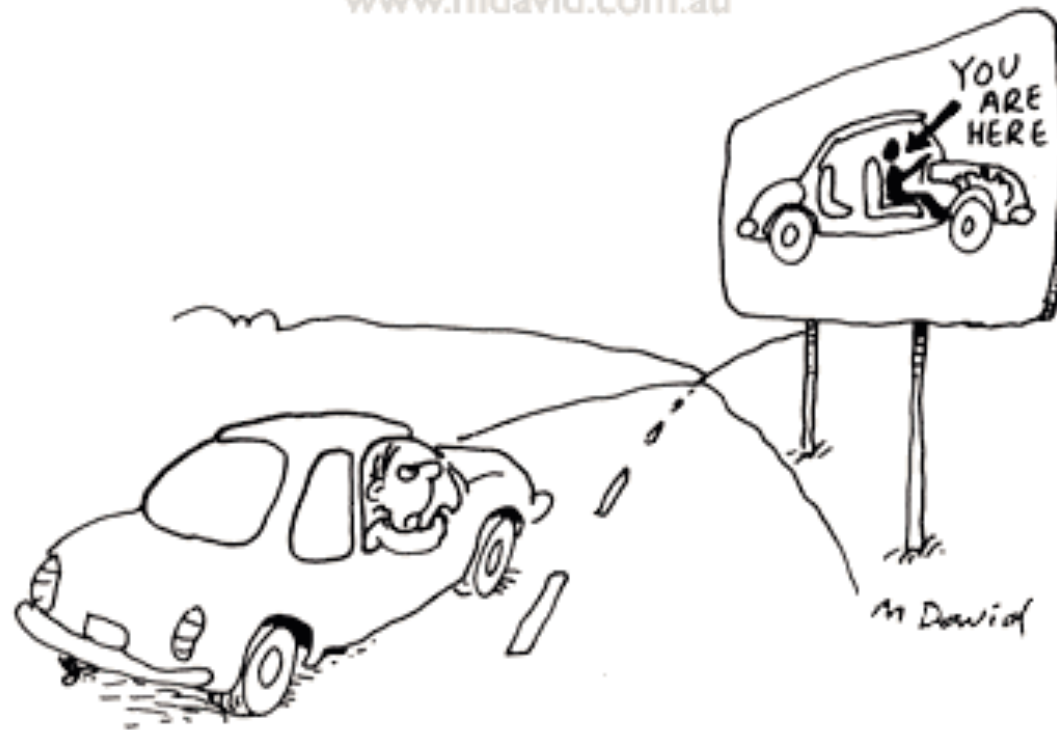
Now

- 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

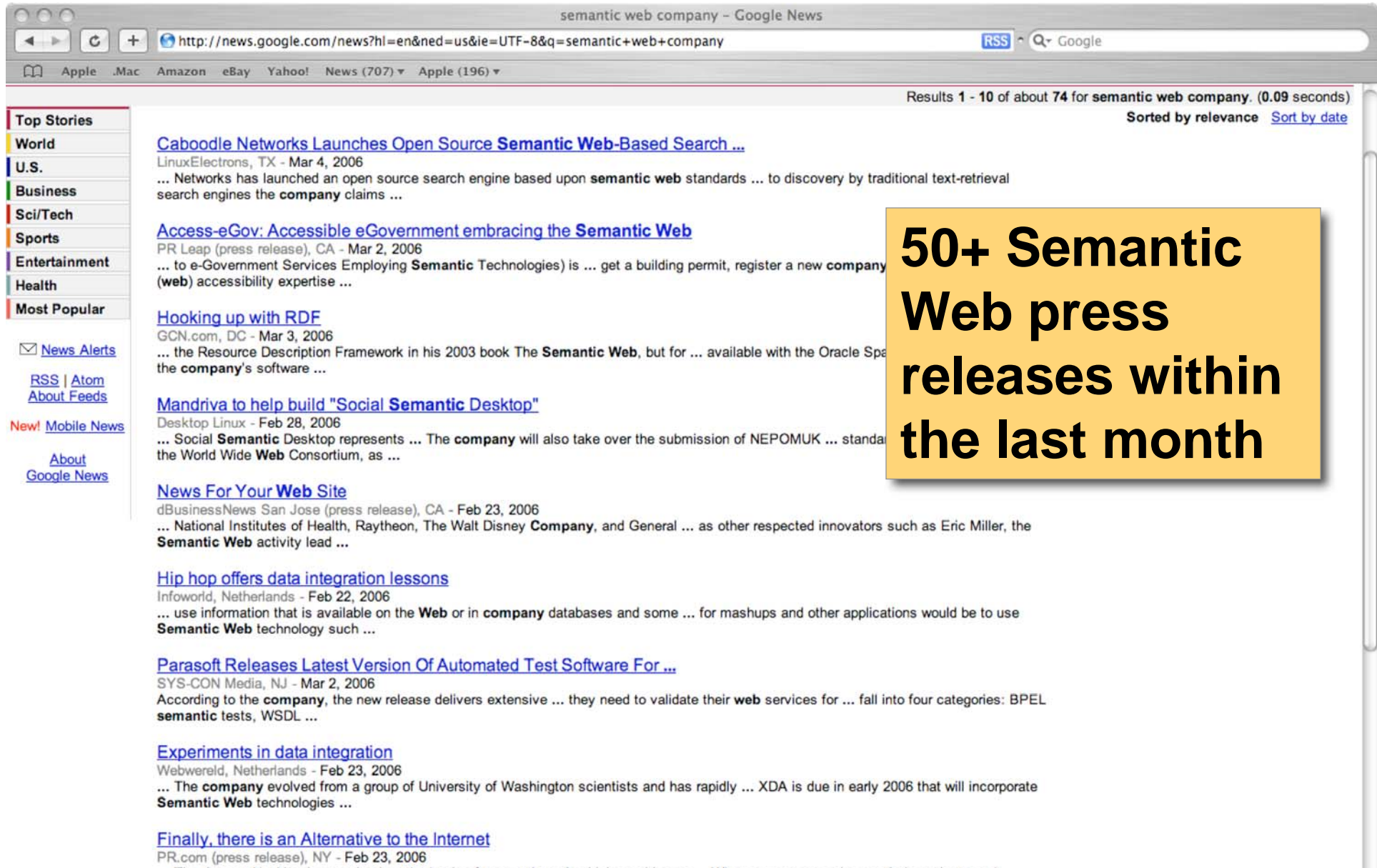


Significant Corporate Activity

- **Semantic (Web) technology companies starting & growing**
 - Cerebra, 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
 - integrator and contractor uptake: Northrop Grumman buys TKS, Lockheed-Martin uses SiberLogic in FCS, SAIC teams with ClarkParsia, ...
 - 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, ...



Significant Corporate Activity



The screenshot shows a Google News search results page for the query "semantic web company". The browser window title is "semantic web company - Google News". The address bar shows the URL "http://news.google.com/news?hl=en&ned=us&ie=UTF-8&q=semantic+web+company". The search results are sorted by relevance, showing results 1-10 of about 74. The results include several articles from early 2006, such as "Caboodle Networks Launches Open Source Semantic Web-Based Search ...", "Access-eGov: Accessible eGovernment embracing the Semantic Web", "Hooking up with RDF", "Mandriva to help build 'Social Semantic Desktop'", "News For Your Web Site", "Hip hop offers data integration lessons", "Parasoft Releases Latest Version Of Automated Test Software For ...", "Experiments in data integration", and "Finally, there is an Alternative to the Internet".

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

Top Stories

- World**
- U.S.**
- Business**
- Sci/Tech**
- Sports**
- Entertainment**
- Health**
- Most Popular**

☒ [News Alerts](#)

[RSS](#) | [Atom](#)
[About Feeds](#)

[New!](#) [Mobile News](#)

[About](#)
[Google News](#)

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
GCN.com, DC - Mar 3, 2006
... the Resource Description Framework in his 2003 book The **Semantic Web**, but for ... available with the Oracle Sp... 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 ... standa... 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 Eric Miller, 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
... The data is ...

50+ Semantic
Web press
releases within
the last month

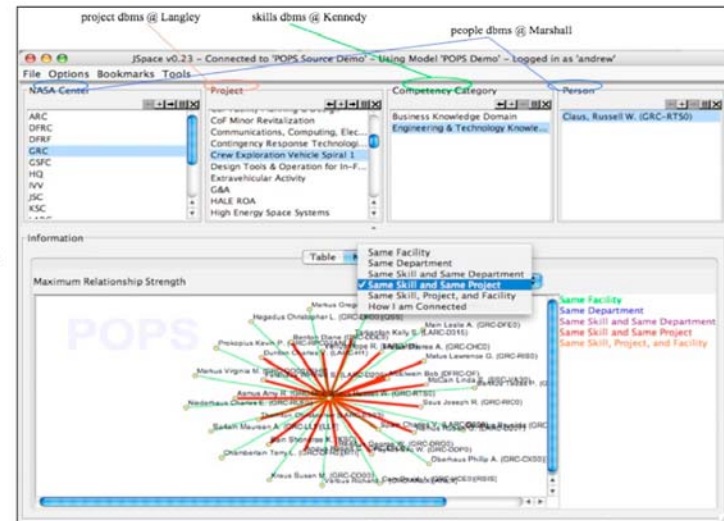
Significant Government Activity

- **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 in 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)

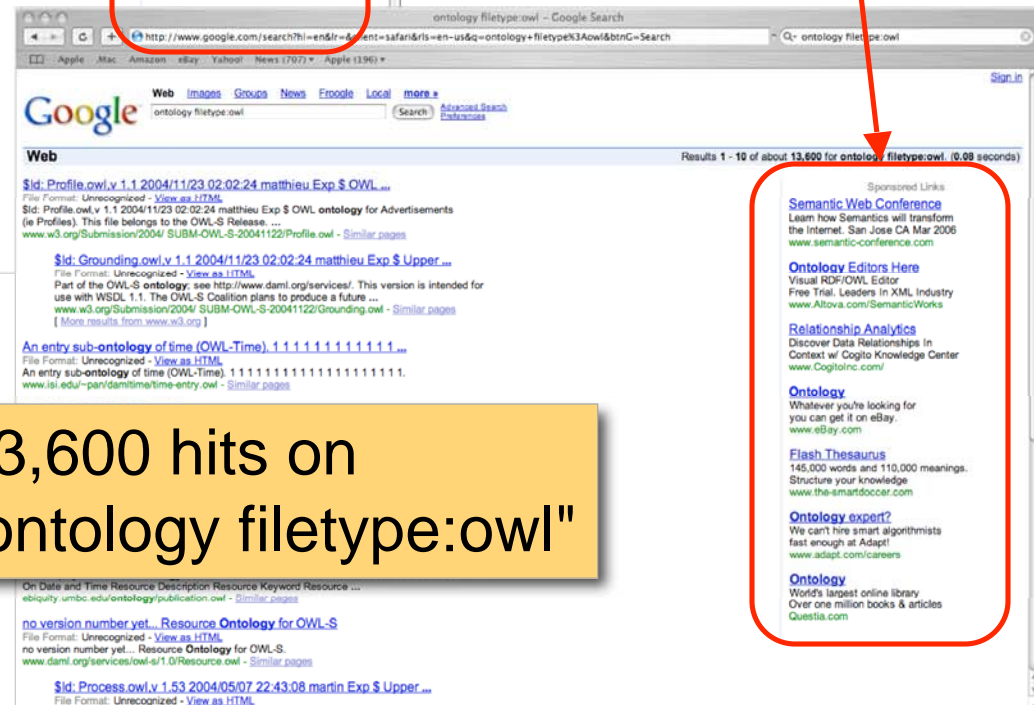


There's a Lot Out There!



Paid ads

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

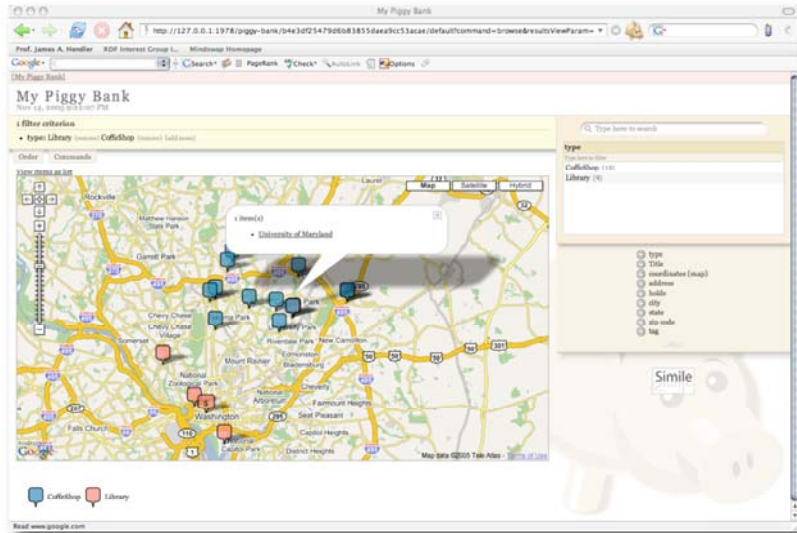


13,600 hits on
"ontology filetype:owl"



(March, 2006)

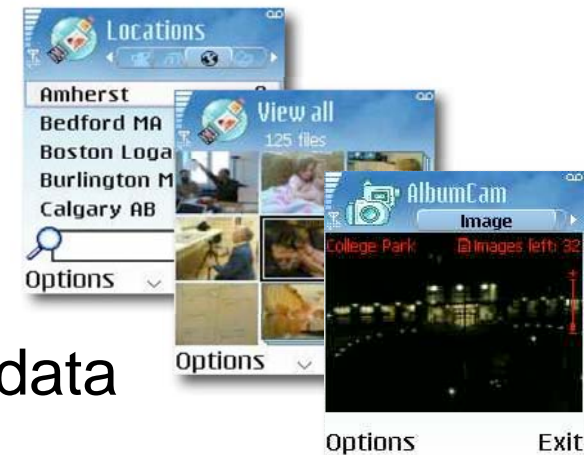
Semantic WEB



Data harvesting &
visualization

A little Semantics
goes a long way

Rich
metadata



Web-based social networks

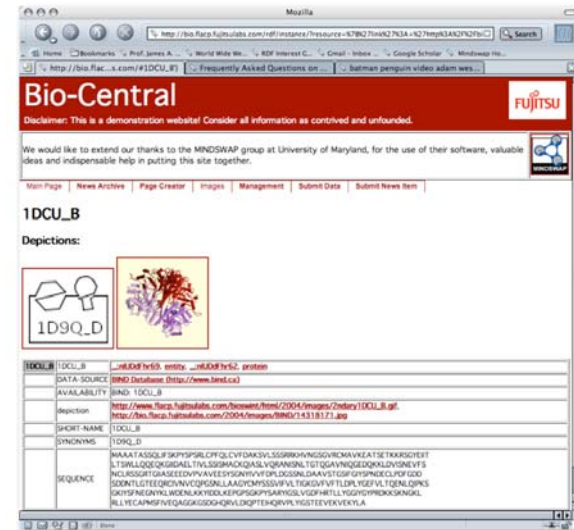
NOKIA
Connecting People



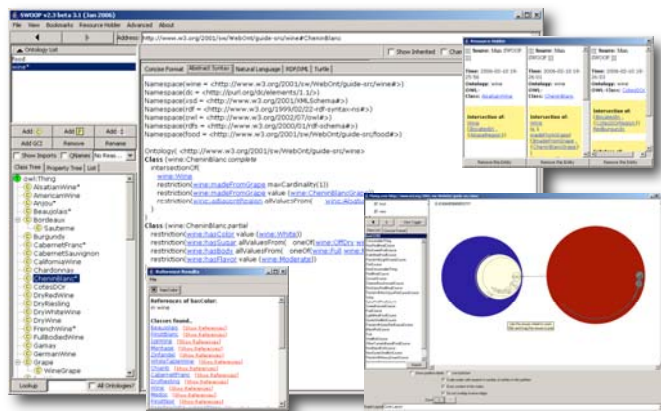
SEMANTIC Web



Digital asset management



Scientific portals

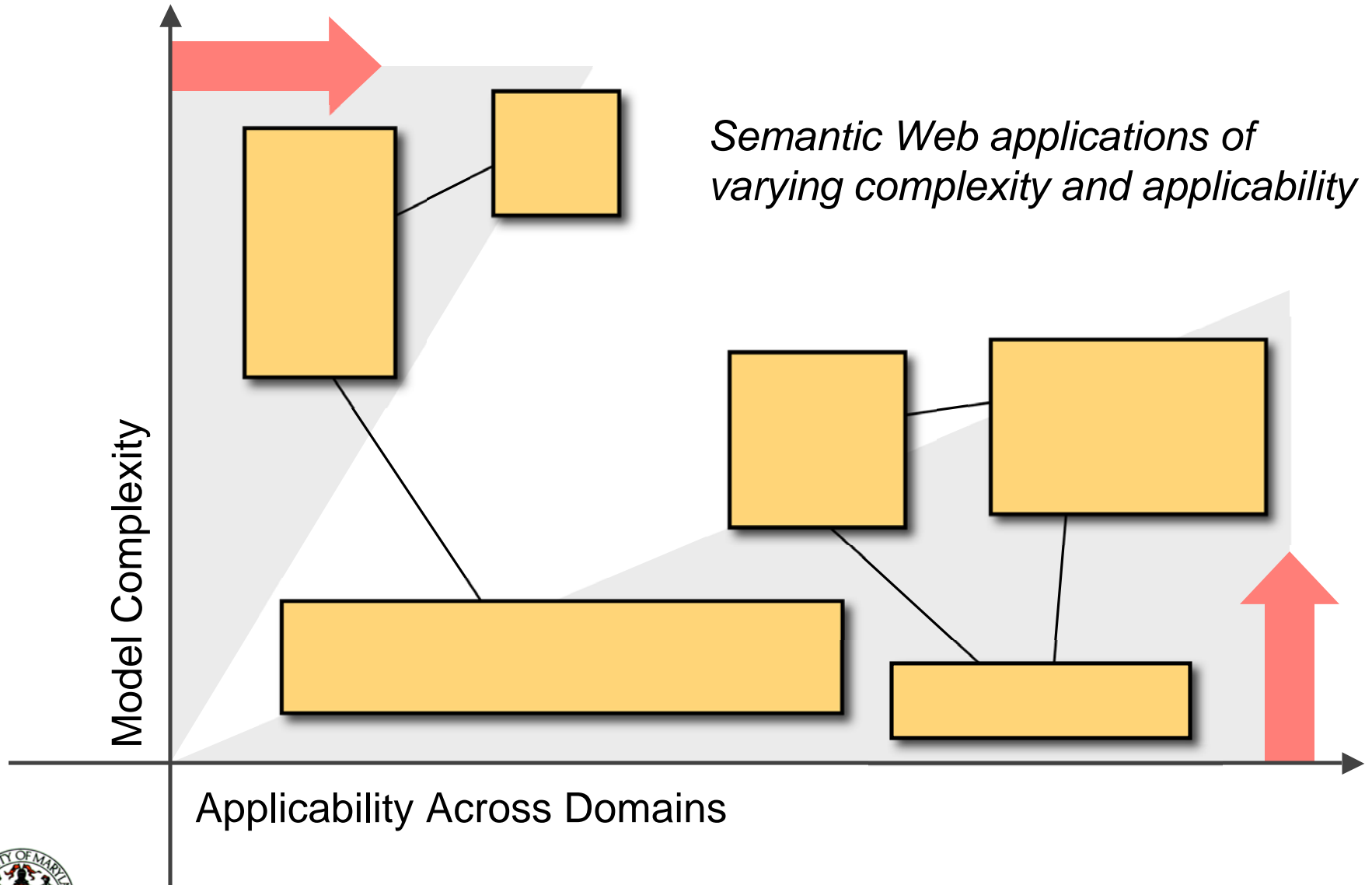


Tools for developers

A little Web
goes a long way

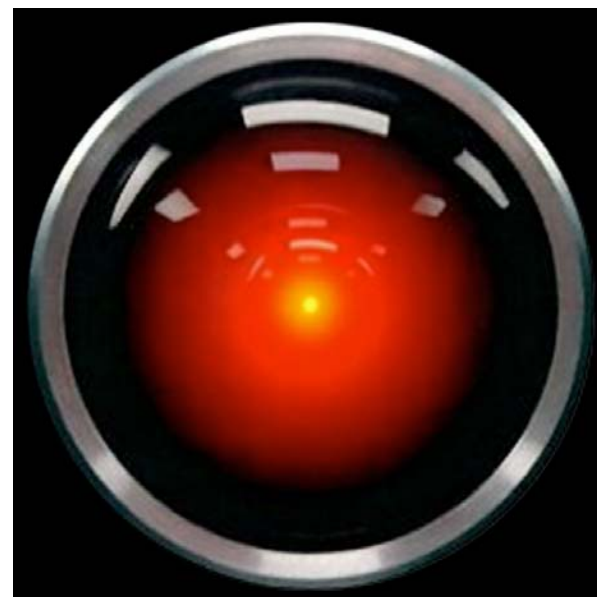


2006: The Gap Is Closing



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!



Semantic WEB Lessons

- **Web needed high value sites**

- personal (homepages, pets)
- public (hobbyists, govt)

- **As these linked up, new functionality emerged**

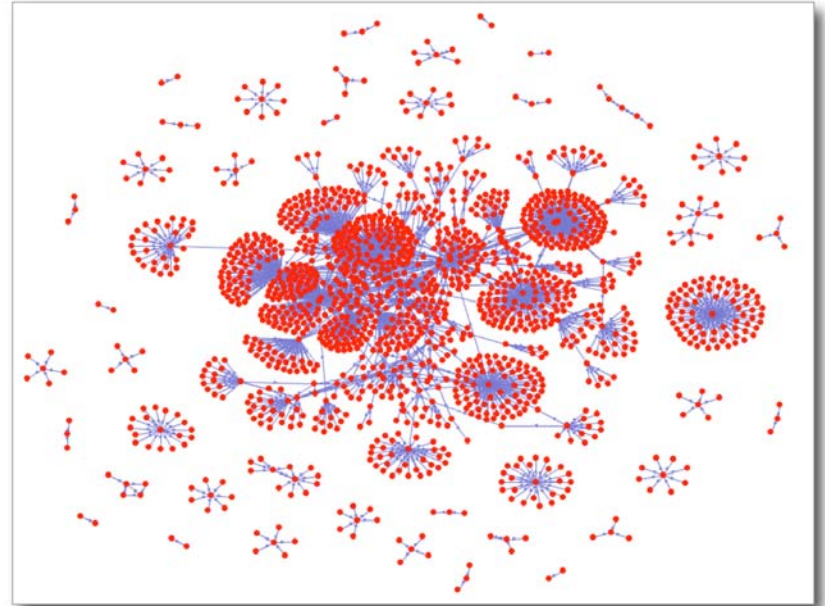
- Yahoo, Alta Vista, ...

- **New business models followed...**

- “give it away” (Netscape)
- marketplace (Amazon)
- advertising (Yahoo, Google)

- **Semantic Web?**

- **SHARE; GIVE IT AWAY!**



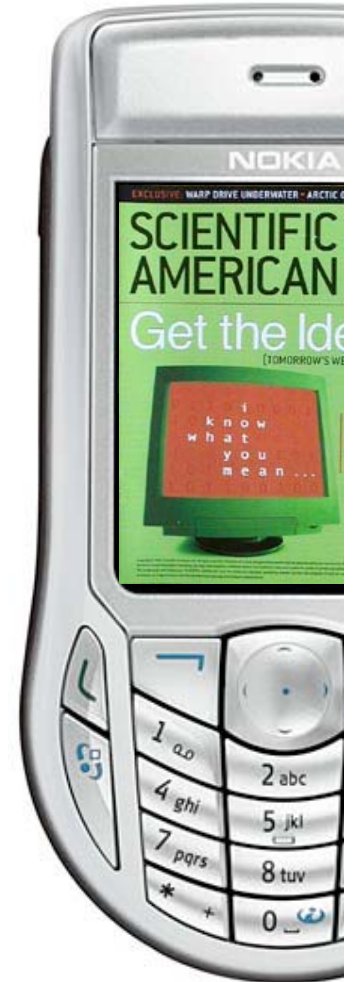
- **What do we need?**

- Open Source Tools
- Open Source Datasets
- Open Source Harvesters

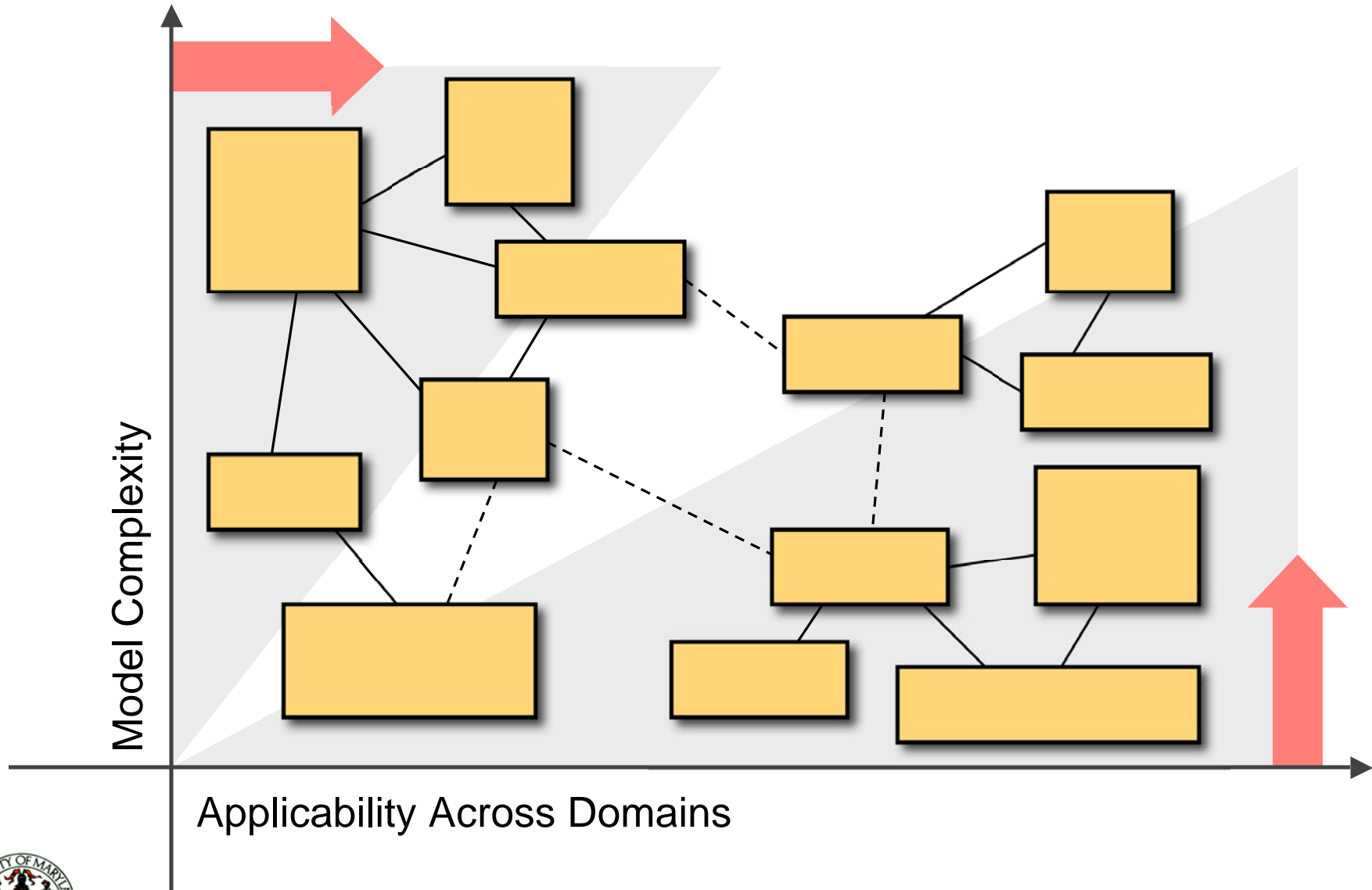


Where Are the Agents?

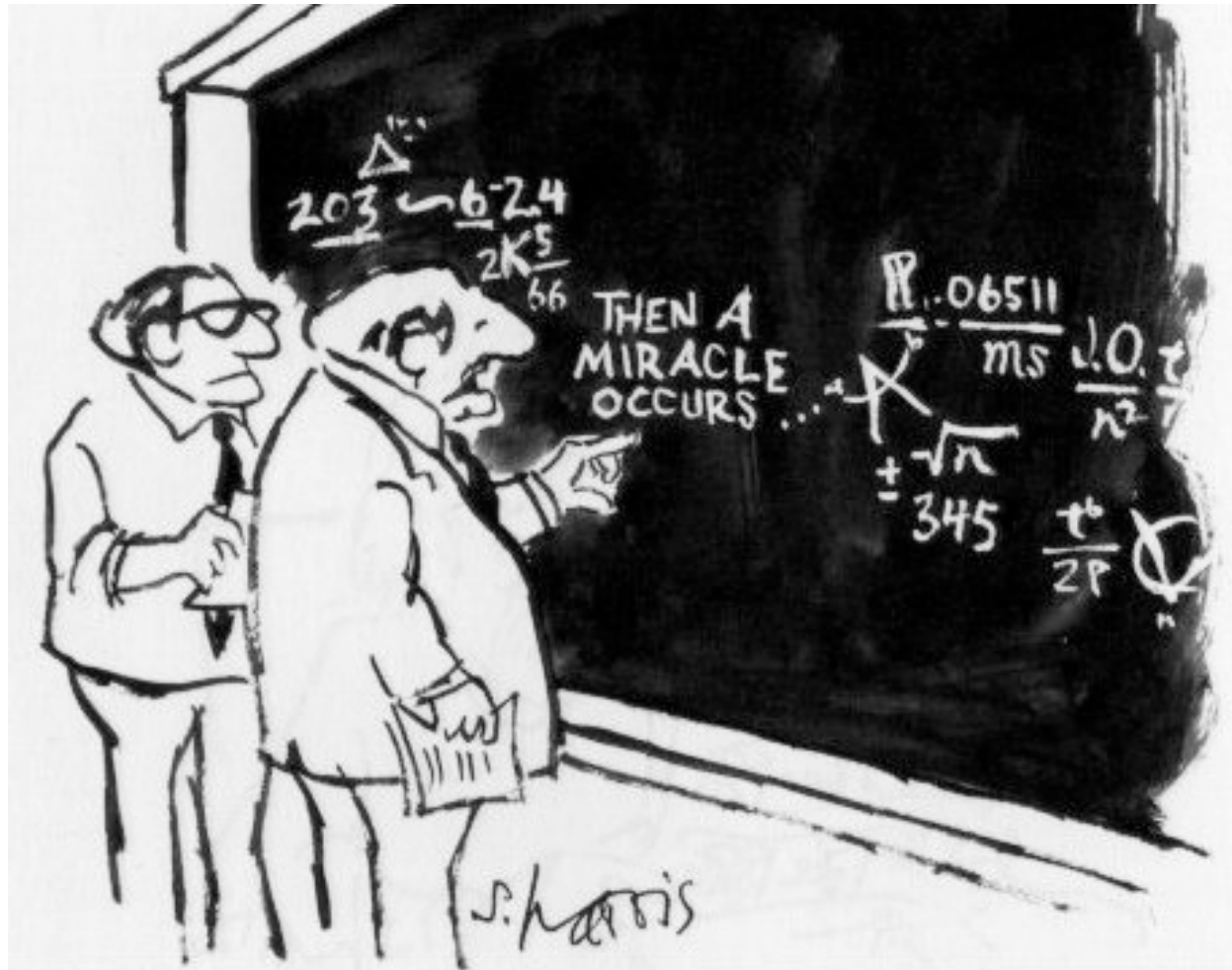
- **“Brave New Applications”**
 - operate autonomously in “unanticipated” situations
 - exhibit robustness in the face of
 - changing, inconsistent and unexpected data
 - variations in reliability, trust
 - capable of serendipitous behavior, opportunism
- **Move from the “tool use” of personal computing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



Linking Is Power!



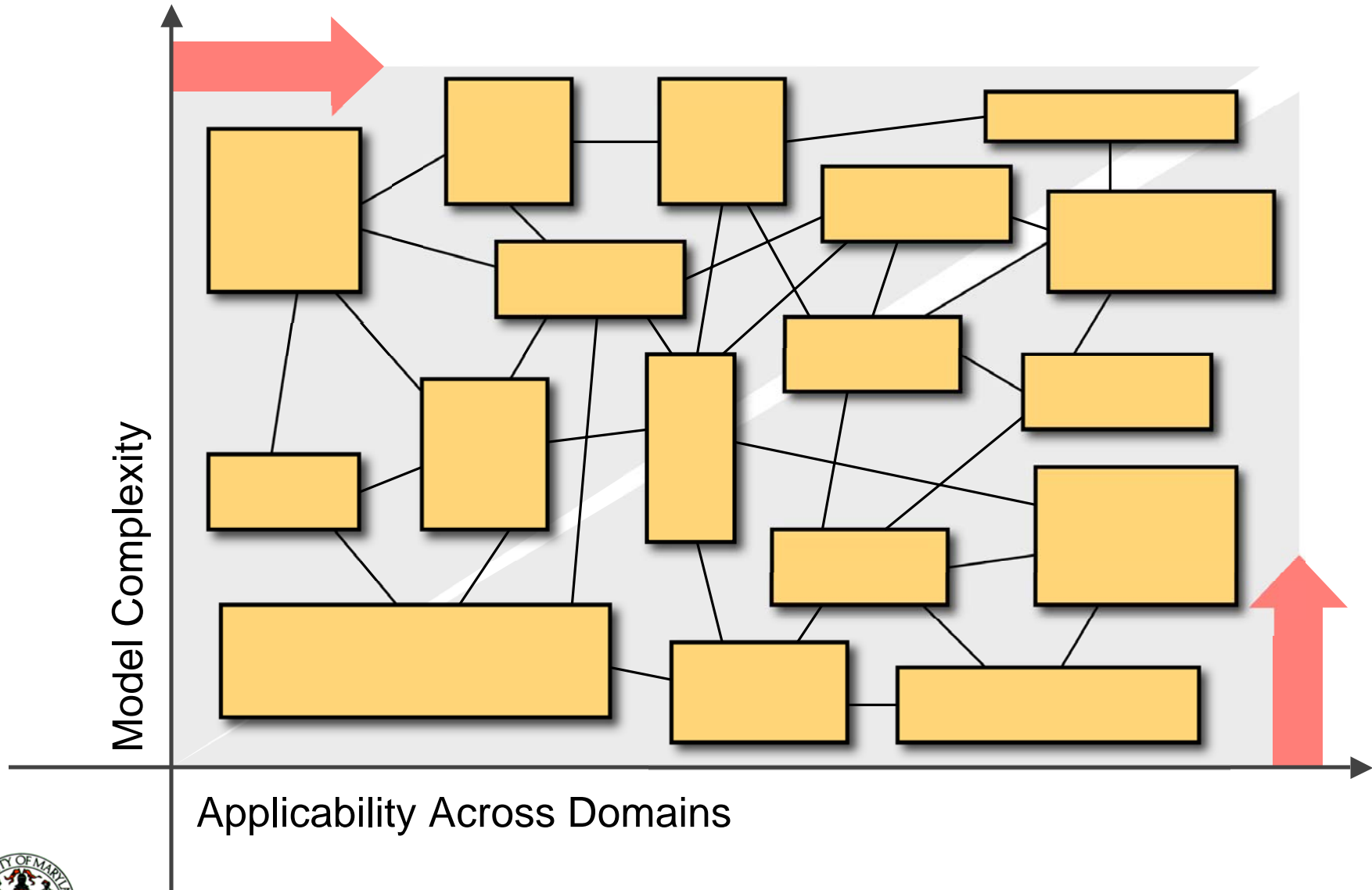
Then a Miracle Occurs...



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

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

Looking Further Out



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?

