

The Semantic Web: Too clever for its own good?

Or: The 7 (f)laws of the Semantic Web



What we'll cover

- Introduction
 - Scope
 - What is the Semantic Web?
 - What difference will it make?
- Semantic Web
 - Past
 - Present
 - The 7 (f)laws of the Semantic Web
 - Future
- Summary

Introduction: Scope of this talk

- Social and Political issues, less technical.
- Critical, with a tiny dash of optimism.

Introduction: What is the Semantic Web?

- The Semantic Web IS NOT:
 - *A replacement for, or evolution of the current Web.*
 - Another attempt at Artificial Intelligence.
 - Semantic -- machines will not be able to understand, or reason from, the data.

Introduction: What is the Semantic Web?

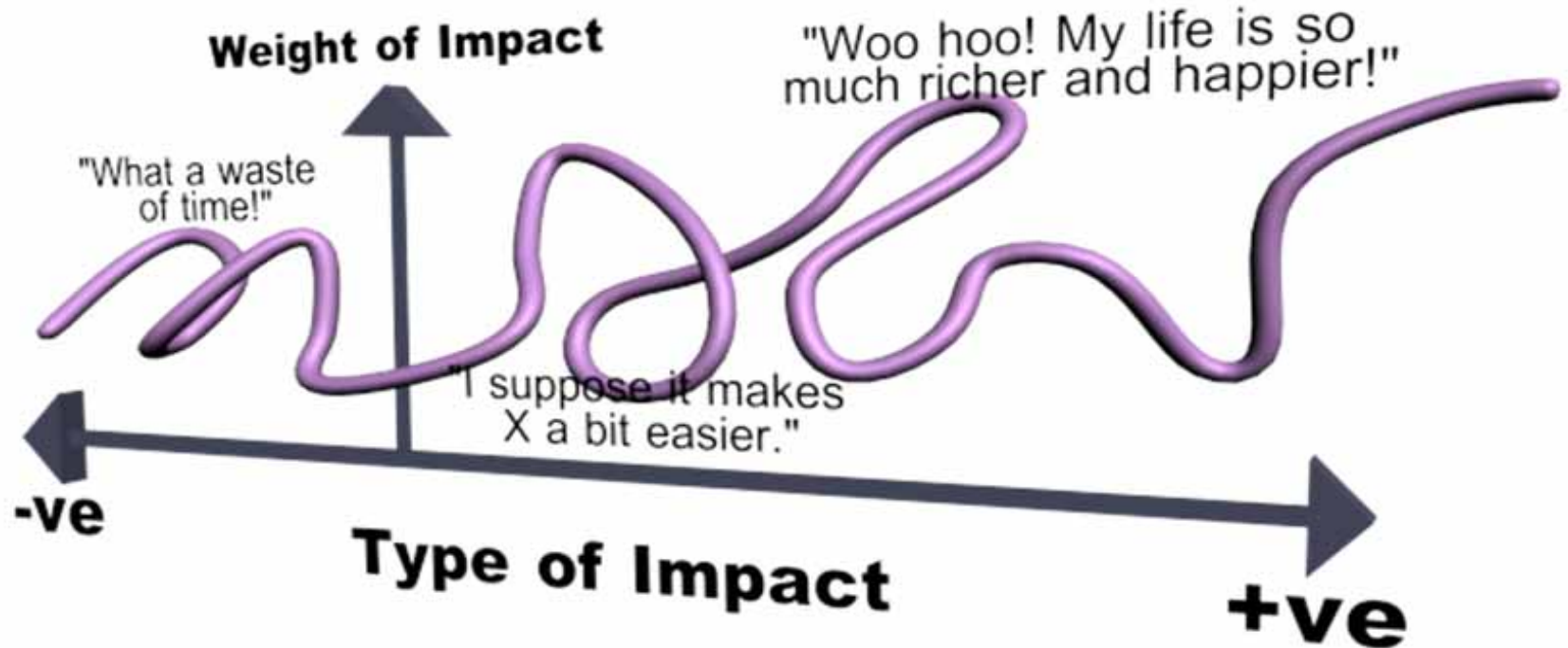
- **The Semantic Web IS:**
 - A web of **linked data** under the current (hyper) **linked resources**.
 - Not built from scratch -- most of this data already exists.
 - But needs to be converted to a new format.

Introduction: What is the Semantic Web?

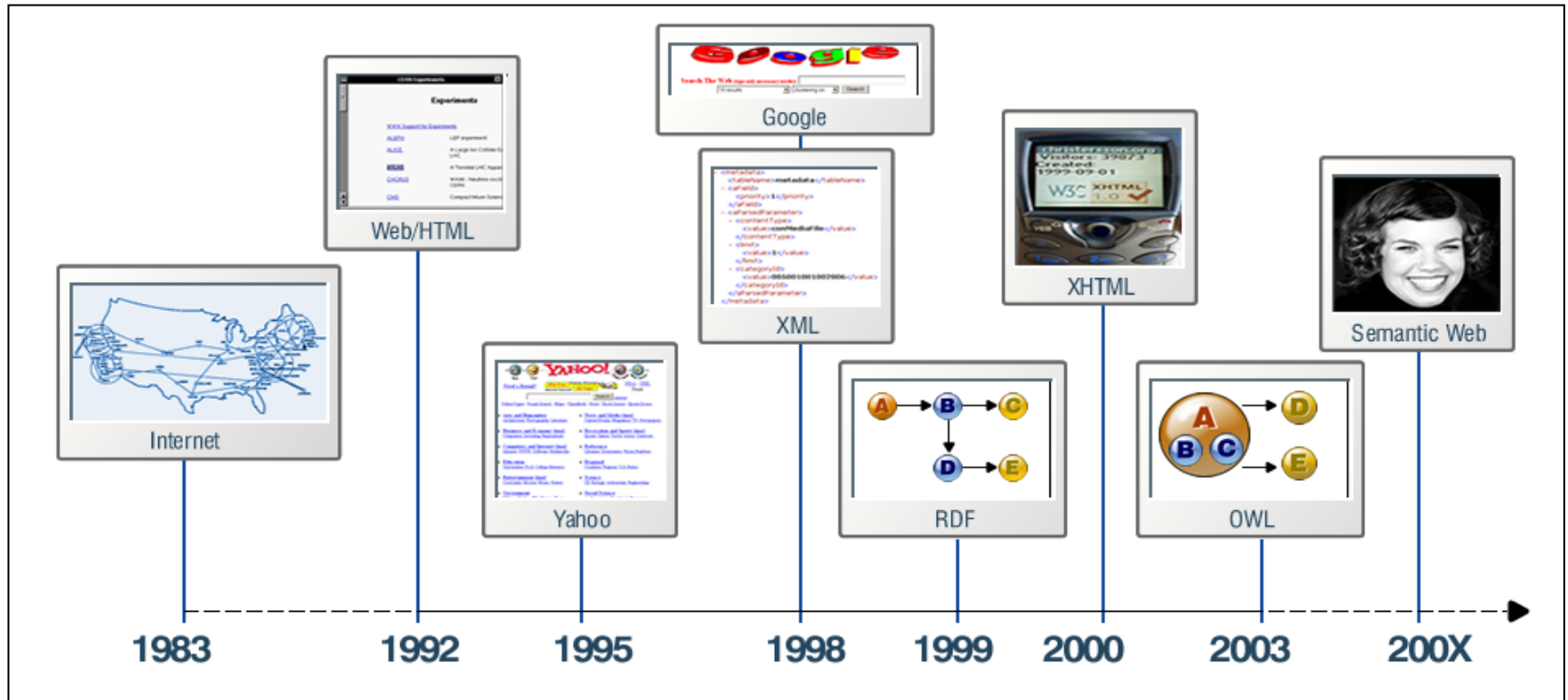
- It might be used for:
 - Finding information/services
 - Metadata
 - An automation infrastructure
 - Annotation
 - Data mining
 - ...

Introduction: What difference will it make?

Your guess is as good as mine.



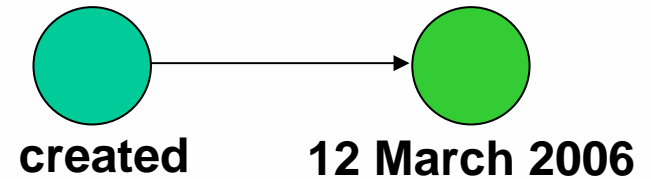
Semantic Web: Past



Semantic Web: Past

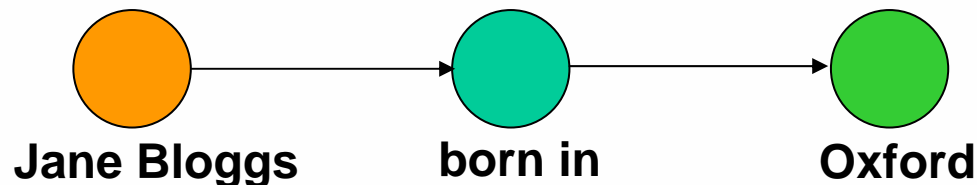
- Metadata (attribute/value pair)

- author = Jane Bloggs
- created = 12 March 2006

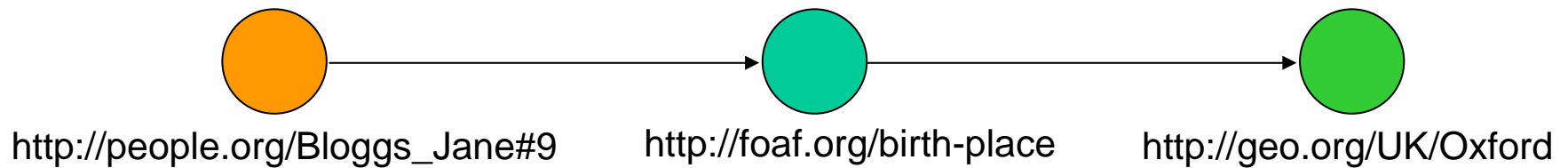


- Metadata (RDF model: 'triple', 'statement', 'graph', ...)

- BCS Homepage -> author -> Jane Bloggs
- BCS Homepage -> created -> 12 March 2006



Semantic Web: Past

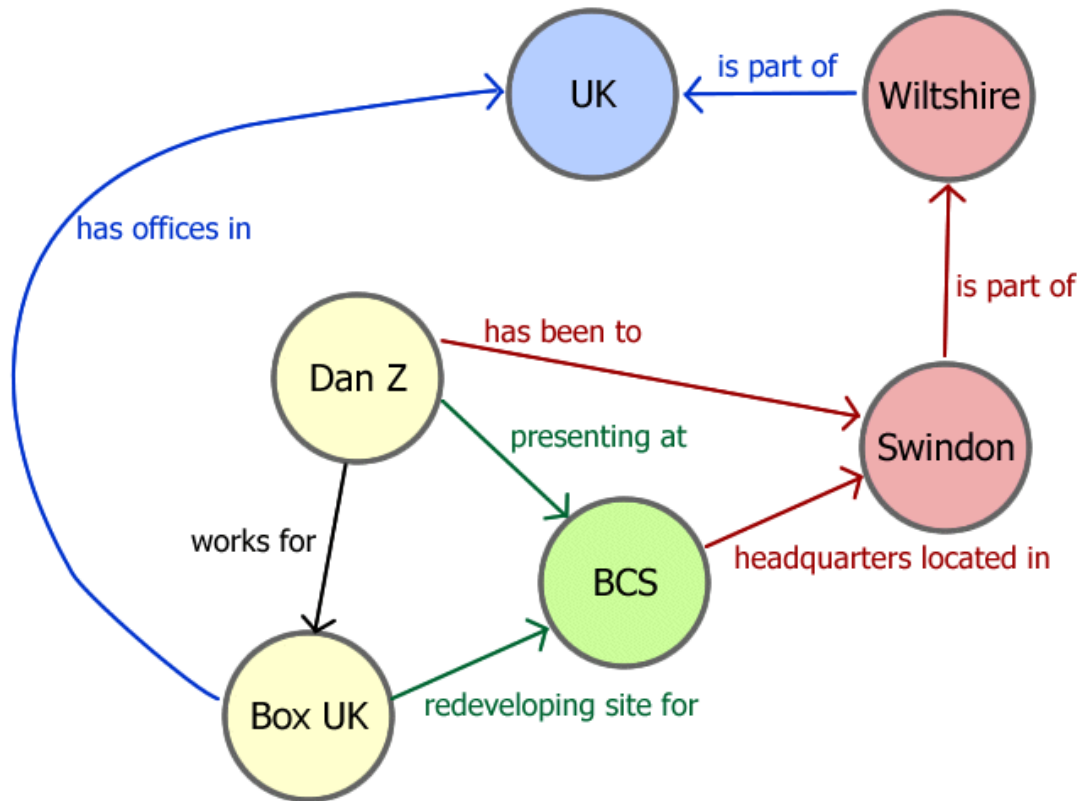


URIs provide standard terms, RDF/XML a standard format.



```
<rdf:Description rdf:about="http://people.org/Bloggs_Jane#9">  
  <foaf:birthPlace rdf:resource="http://geo.org/UK/Oxford" />  
  <foaf:birthDate>1965-09-25</foaf:birthDate>  
</rdf:Description>
```

Semantic Web: Past

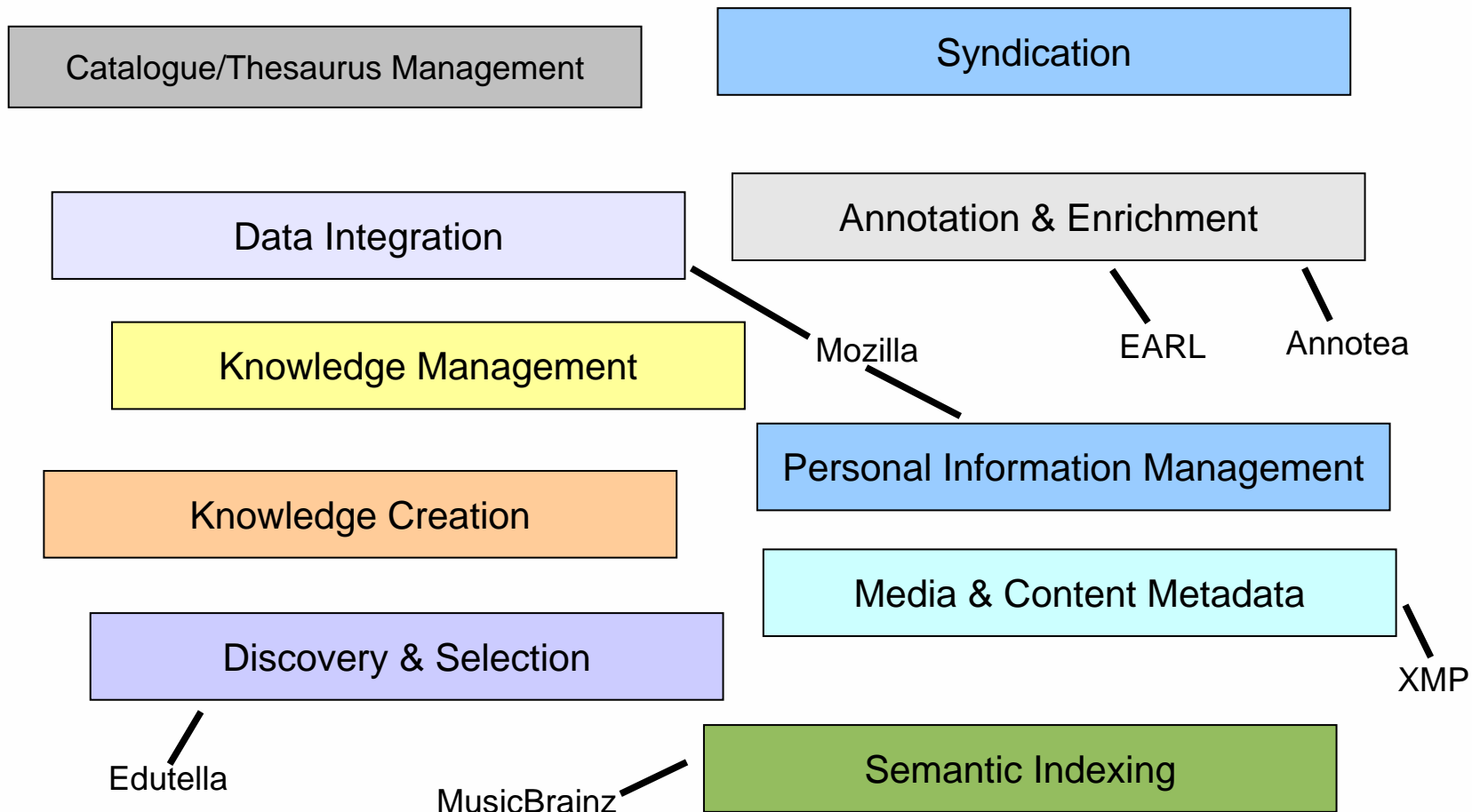


Semantic Web: Past

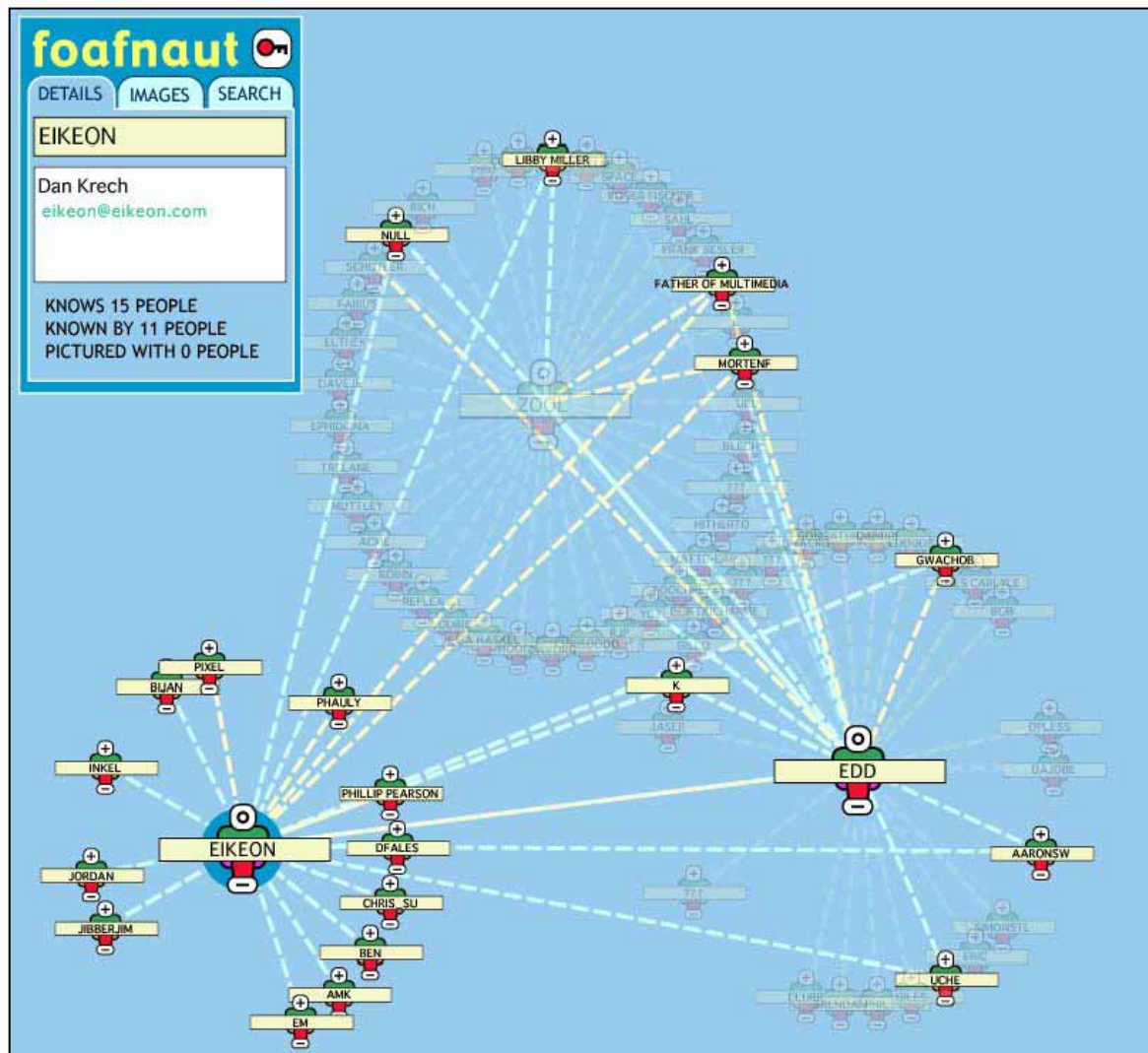
- **RDF Schema** - basic vocabularies:
 - continent
 - country
 - county
 - » city
 - » town
 - relative
 - sibling
 - brother
 - sister
- **OWL** - complex ontological concepts, including:
 - Cardinality
 - Transitivity
 - Equivalence
 - Unions, Intersections and Disjoints

Semantic Web: Present

- 100+ Semantic Web applications, covering:



Semantic Web: Present



Semantic Web: Present

YAHOO! SEARCH

Creative Commons Search **BETA**

Search CC

Search the Web

- Find content I can use for commercial purposes.
- Find content I can modify, adapt, or build upon.



[Advanced Search](#)
[My Web](#) **BETA**

Why is this search different?

This Yahoo! Search service finds content across the Web that has a [Creative Commons](#) license. While most stuff you find on the web has a full copyright, this search helps you find content published by authors who want you to share or reuse it, under certain conditions. [Learn more...](#)

Forums & Feedback

See what others are saying about Yahoo! Creative Commons search in [forums](#), or send us [feedback](#).

Copyright © 2005 Yahoo! Inc. All rights reserved. [Privacy Policy](#) - [Terms of Service](#) - [Copyright/IP Policy](#) - [Submit Your Site](#) - [Job Openings](#)

Semantic Web: Present

The screenshot displays the Haystack web application interface. At the top, there is a navigation bar with a search field and several menu items: Back, Go to, Search for, Starting Points, Information Sources, and Buddy List. The main content area is titled "Home" and includes a sub-header "Show with collection elements" and "Lay out like a portal".

On the left side, there are three panels:

- Scrapbook:** Contains a link to "Haystack Home Page".
- Active Tasks:** Lists several tasks with checkboxes and arrows, such as "Ask Dave about financing", "Buy Mother's Day gift", "Call electric company", "Go grocery shopping", "HIV 2003", and "Reconfirm flight".
- HIV 2003:** A collection panel showing "Items in this collection" and "No items in list".

The central area features an **Inbox** and a **Weather Agent**.

The **Inbox** shows a list of emails:

- John Doe: Wed May 14, 2003, 11:31 AM. Subject: Flight info.
- Mary Smith: Wed May 14, 2003, 10:15 AM. Subject: Cool paper at HIV 2003.
- Ippanno: Wed May 14, 2003, 9:27 AM. Subject: International House of Sushi. Body: For sushi lovers.
- John Doe: Sat March 1, 2003, 4:31 PM. Subject: Draft announcement.
- John Doe: Sun January 19, 2003, 4:31 PM. Subject: Murine cydin T1.
- John Doe: Sun January 19, 2003, 4:31 PM.

The selected email from John Doe is displayed in detail:

- From:** John Doe available for chat
- Sent:** Wed May 14 10:31:41 EST 2003
- To:** Dennis Quan
- Cc:** No items in list
- Subject:** Flight info
- In Reply To:** No items in list
- Attachment:** Flight to San Francisco, Impedia.net
- Body:** Hi Dennis, I've attached my flight info. Also, you may want to try out this new travel agency; it's really easy to use. John

The **Weather Agent** shows the latest forecast for Cambridge, MA (02139):

- Last updated 10 PM EDT 18 JUN 03
- Sunrise 5:07 AM; sunset 8:24 PM
- overnight drizzle or light rain likely with patchy

On the right side, there are several toolbars and panels:

- Apply tools to:** Flight info
- Add to collection:** Add
- Annotations:** Add
- Organize:** Choose a categorization scheme: Default Categories, HIV 2003 stuff, Work categories. Add
- Categorize Flight info:** Cardiac muscle (2 items), Collaboration (3 items), HCI (3 items), March 14 AcmeMail (2 items) talk. Browse full screen, Add
- Recommended operations:** Forward, Mark as read, Mark as unread, Remind me to reply to this, Reply, View discussion.
- Related activities:** Add

At the bottom, a status bar indicates: "User Dennis is logged into Haystack" and "Resolving LSID <urn:lsid:ncbi.nlm.nih.gov:lsid.i3c.org:pubmed:12458222>".

Semantic Web: Present

Combined Information: restaurants + movies - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://127.0.0.1:1978/piggy-bank/model1116627657351?command=

[My Piggy Bank] | Combined Information

restaurants + movies


Combined Information

1 filter criterion

- type: Show (remove) Restaurant (remove)

Order Commands

View items as list



Map - Satellite New!

1 item(s)

- ◆ [Abacus Cuisine Of China](#)

©2005 Google

Map data ©2005 NAVTEQ™, Tele Atlas

Show Restaurant

Transferring data from mt.google.com...

Type here to search

cuisine

Type here to filter

- "Asian" (28)
- "Californian" (1)
- "Canadian" (2)
- "Chinese" (5)
- "Desserts" (1)
- "Dim Sum" (4)

date

By hour

Type here to filter

- 10:00 PM (138)
- 11:00 PM (20)

type

Type here to filter

- Restaurant (28)
- Show (158)

address

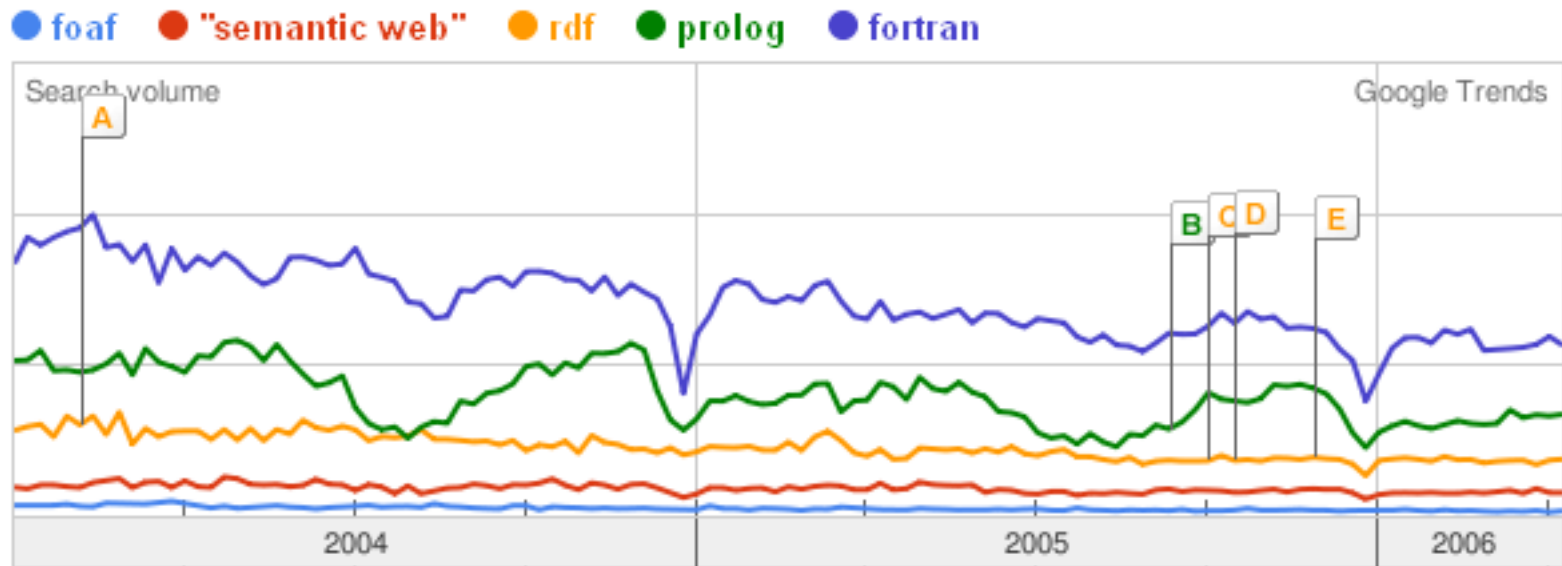
Type here to filter

Semantic Web: Present

- So how much RDF data is on the web?
 - About **4 million** RDF files
 - **FOAF**: About 300,000
 - **RSS**: Most of the remaining

(Ballpark estimates, using google.com)

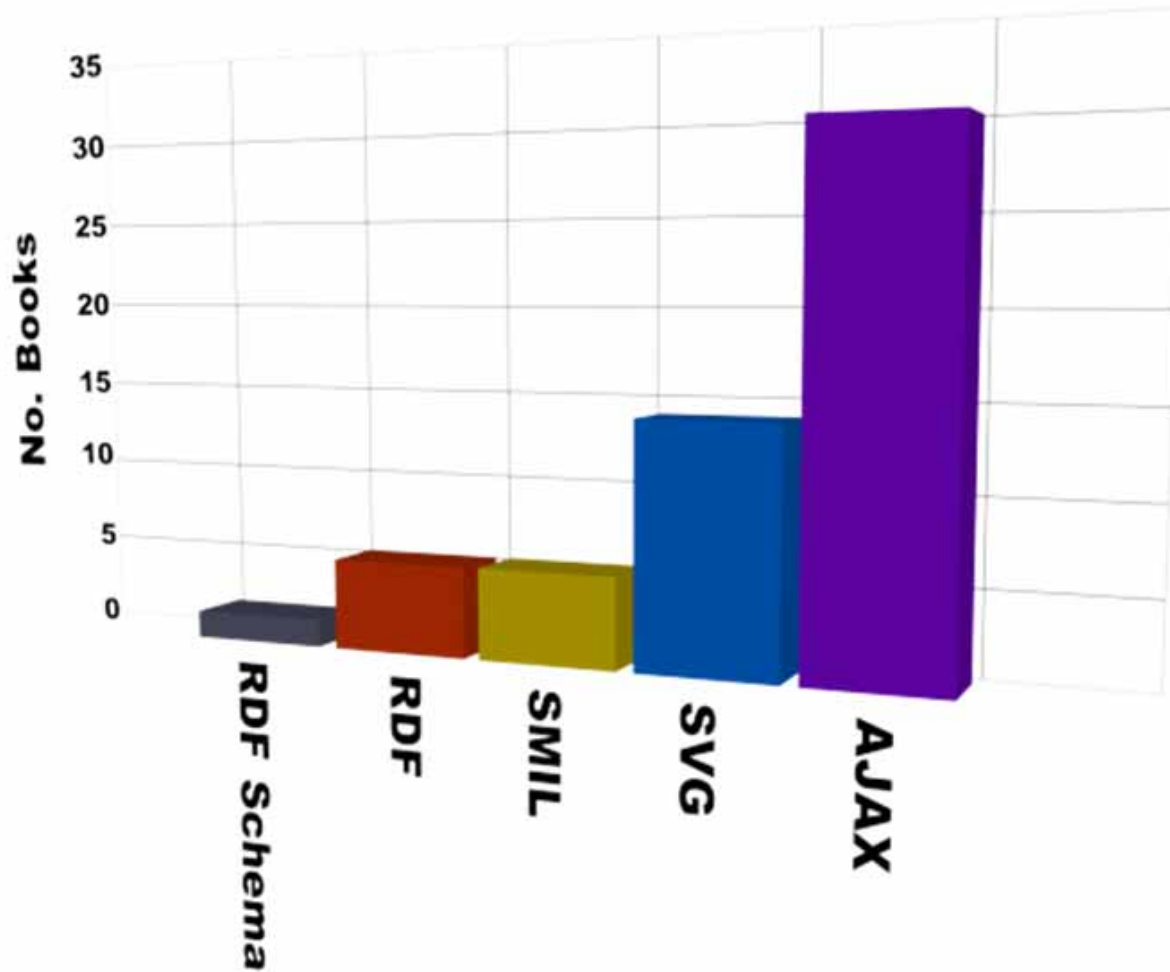
Semantic Web: Present



Interest in the Semantic Web flat-lined a long time ago.

(source: google.com/trends)

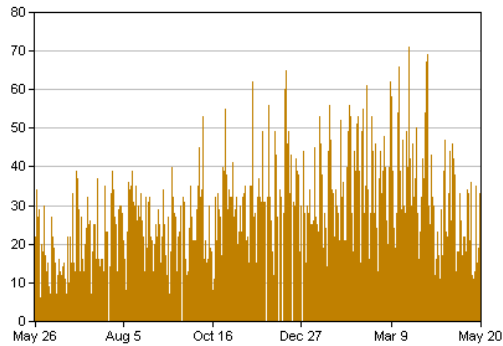
Semantic Web: Present



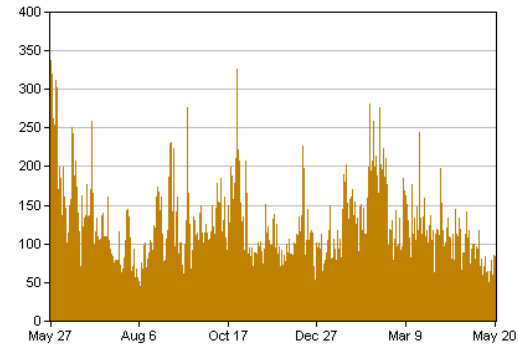
Books in Amazon.com under *Computers & Internet* category.

Semantic Web: Present

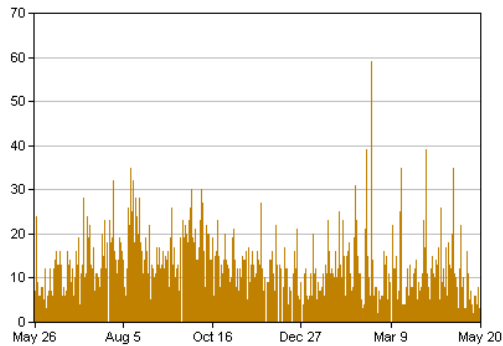
Semantic Web (Avg ~20)



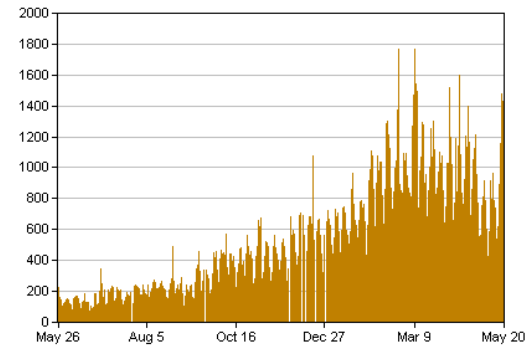
RDF (Avg ~100)



FOAF (Avg ~10)



AJAX (Avg ~500)



Blogs by subject, per day, for last year (source: technorati.com)

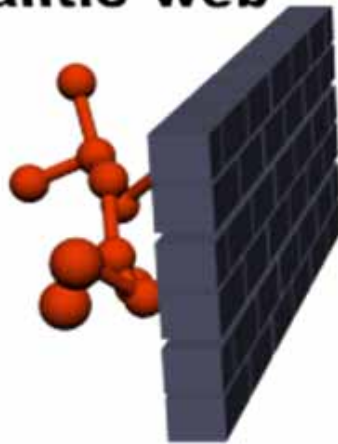
The 7 (f)laws of the Semantic Web.

The 1st (f)law of the Semantic Web

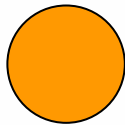
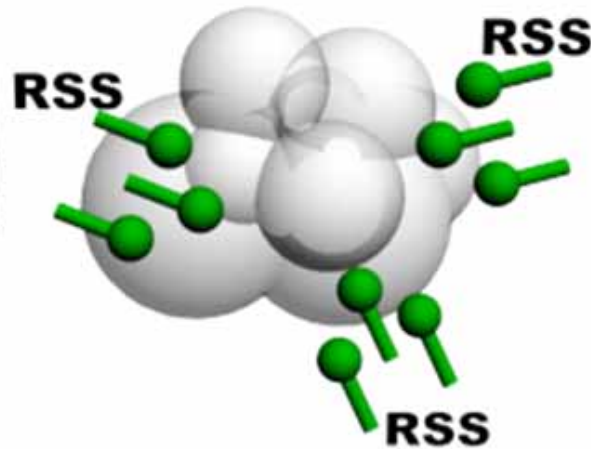
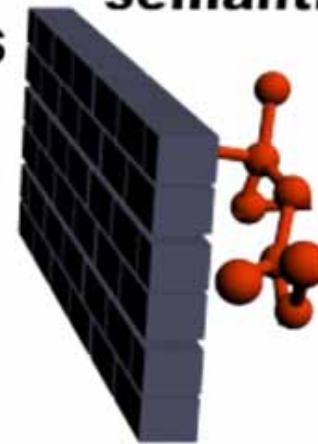
Not all Semantic Web
data are created equal.

The 1st (f)law of the Semantic Web

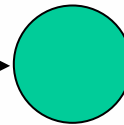
**Life Sciences
"semantic web"**



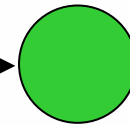
**Crime & Police
"semantic web"**



<http://bbc.co.uk/news-12345>



<http://rss.org/subject>



"Science/Nature"

The 2nd (f)law of the Semantic Web

A technology is only as good as developers think it is.

The 2nd (f)law of the Semantic Web

- Most successful technologies can be learnt from example files. RDF/XML is a bit trickier.
- The Semantic Web community is (or appears to be) academic, inward-looking and uninviting.

The 2nd (f)law of the Semantic Web

“IMHO, RDF/RDFS/OWL are not well suited to proving validity, due to the open world assumption. There are usually too many possibilities to prevent any incorrect interpretations.

On the other hand, it does just fine with consistency. The only trick is that people are often surprised that many constructs can be considered consistent... again due to the open world assumption. As an example, I first learnt this when I discovered that (under OWL) a predicate with cardinality of 1 for class C can be used multiple times on a single instance of C.”

(A recent post from the “general interest” W3C Semantic Web mailing list)

The 2nd (f)law of the Semantic Web

- "I've developed an overall impression of dismay at the latest RDF model semantics specs."
- "I get the feeling that in trying to achieve the ontological purity needed for the Semantic Web, it's starting to leave the desperate hacker behind."

(A respected developer and blogger)

The 2nd (f)law of the Semantic Web

- "I've lost a lot of my interest in working with the Semantic Web lately, and I don't see it coming back anytime in the near future."
- "...frustration with evangelizing being part of the process of proceeding in the Semantic Web world. Every time I take a step forward with some code, I find another 5 steps I have to take back in order to defend my position and the way I've done it."

(A Semantic Web developer and blogger)

The 2nd (f)law of the Semantic Web

- In response to these complaints:

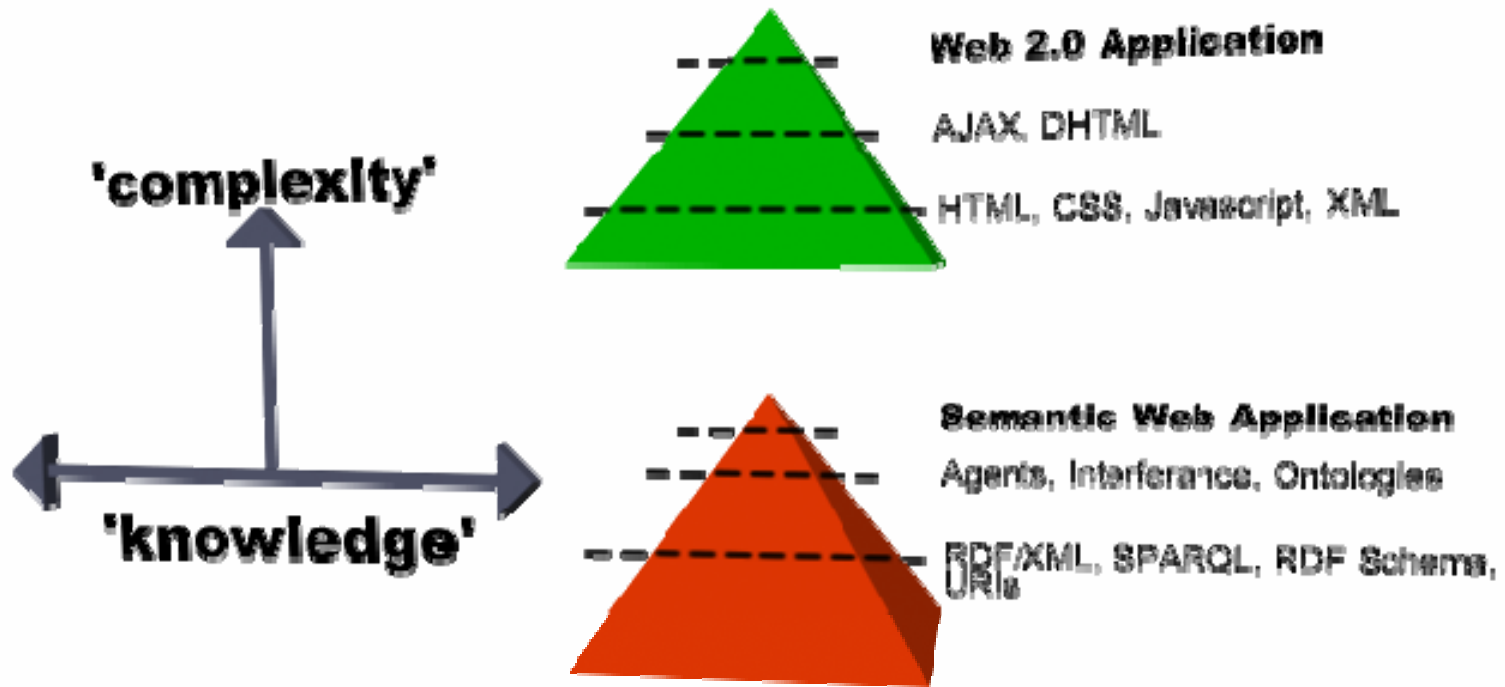
"Time to stop apologizing for the spec. Time to stop bemoaning the complex bits. Don't like it, don't use it."

[Author of one of the mainstream RDF books]

The 3rd (f)law of the Semantic Web

Complex Systems must
be built from successively
simpler systems.

The 3rd (f)law of the Semantic Web



The 4th (f)law of the Semantic Web

A new solution should
stop an obvious pain.

The 4th (f)law of the Semantic Web

- Let's take another look at what the Semantic Web could provide:
 - Finding information/services
 - Metadata
 - An automation infrastructure
 - Annotation
 - Data mining

The 5th (f)law of the Semantic Web

People aren't perfect.

The 5th (f)law of the Semantic Web

- Creating metadata and classifications is difficult.
- People are biased and fallible.
- People have different contexts, different needs.

The 6th (f)law of the Semantic Web

You don't need an
Ontology of Everything.

But it would help.

The 7th (f)law of the Semantic Web

Philanthropy isn't
commercially viable.

Semantic Web: Future

- W3C currently have specifications for:
 - RDF Model
 - RDF/XML
 - RDF Schema
 - OWL
- And are working on ('Phase 2'):
 - Semantic Web for Multimedia Content
 - RIF (Rule Interchange Format)
 - SPARQL
 - Semantic Web Services
 - Trust Policies
 - ...

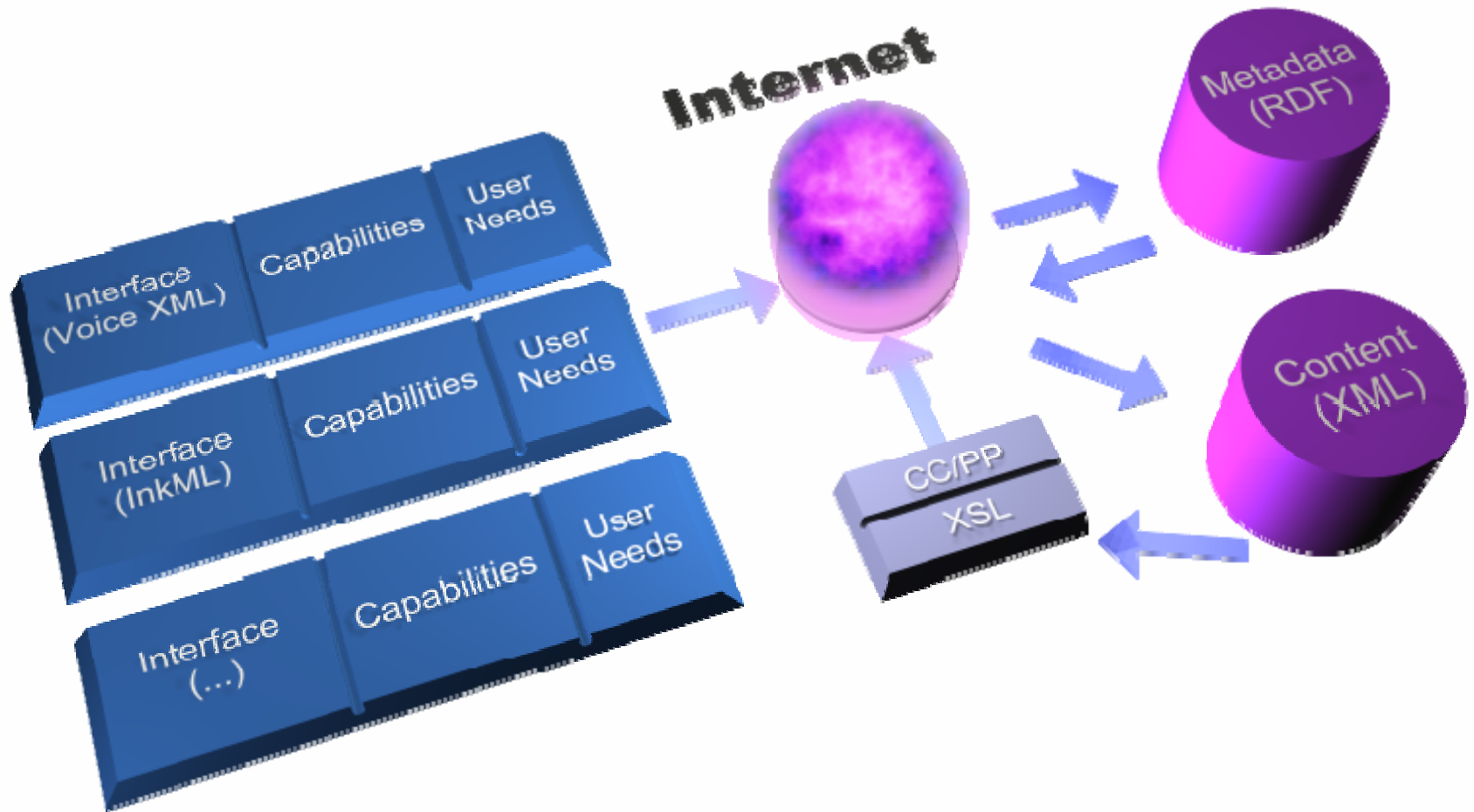
" ... I have seen *lots* of total crap work that was based on specs. It's *the* single worst way to write software, because it by definition means that the software was written to match theory, not reality."

Linus Torvalds

Semantic Web: Future

- Opportunities:
 - Microsoft could help build the Semantic Web developer community.
 - Google could release a Semantic Web search.
 - Or maybe the underdogs/competitors:
 - Adobe
 - Yahoo!
- Threats:
 - Semantic Web specifications collapse under their own weight.
 - Simpler, alternative solutions (microformats, 'tagging', Google Base).
 - Resolution of 'Trust' issue.

Semantic Web: Future



Semantic Web: Future

- What you can do:
 - Expose your data as RDF
 - If you need help, ask – make it clear that you WANT to participate.
 - Educate and inform on what the Semantic Web is/could be.

Summary/Conclusions

- Huge potential, huge - but not yet critical - problems.
- The basic technologies are ready.
 - **The good news:** You just have to save your data in a different format.
 - **The bad news:** It's a pretty complicated format.
- And you probably won't see any effect. Not for a while, anyway.
- But it could be amazing!

Summary/Conclusions

dan@boxuk.com

www.boxuk.com

- www.foaf-project.org/
- search.yahoo.com/cc
- haystack.lcs.mit.edu/
- simile.mit.edu/piggy-bank/

