The Semantic Web:

Too clever for its own good?

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

Potentially
Powerful

But... can't
cope with stairs.

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

- Social and Political issues, less technical.

- Critical, with a tiny dash of optimism.

The Semantic Web IS NOT:

- A *replacement for*, or *evolution of* the current Web.

- Another attempt at Artificial Intelligence.

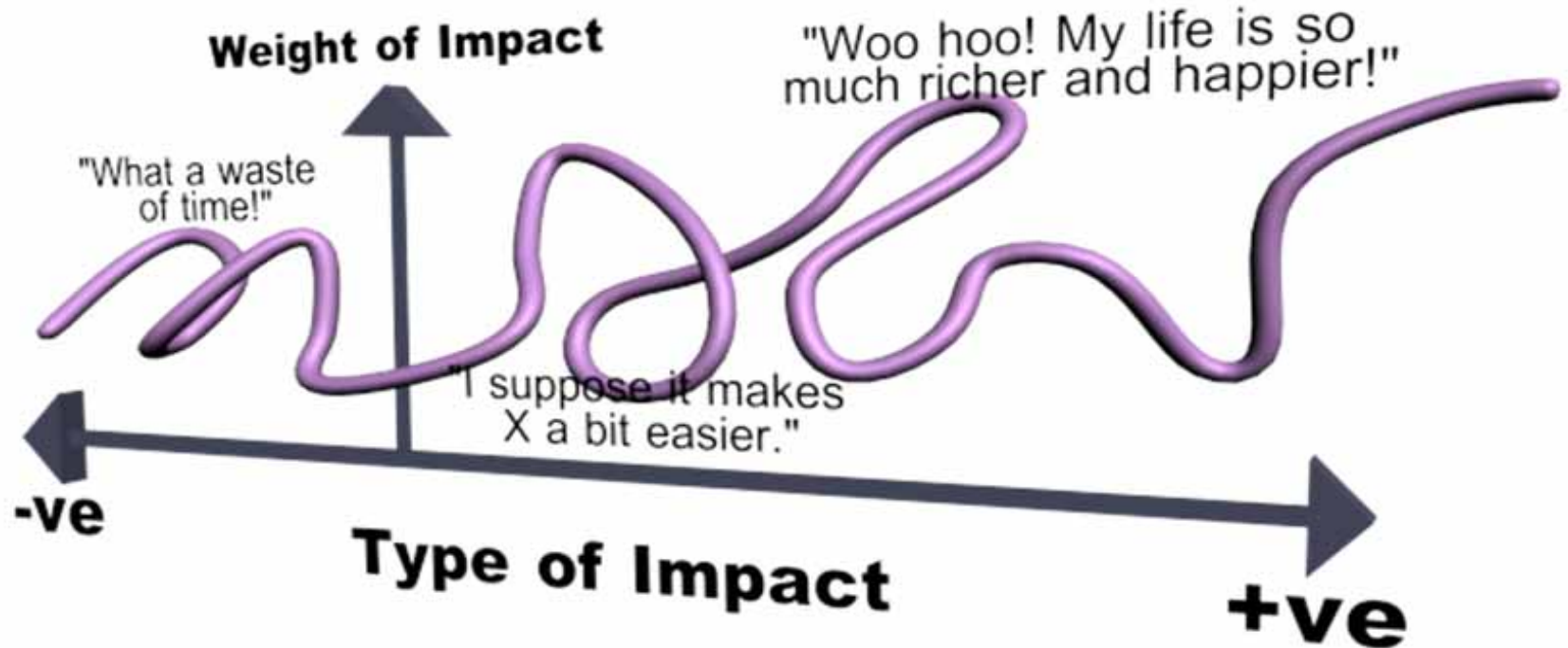- Semantic -- machines will not be able to understand, or reason from, the data.

■ 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.
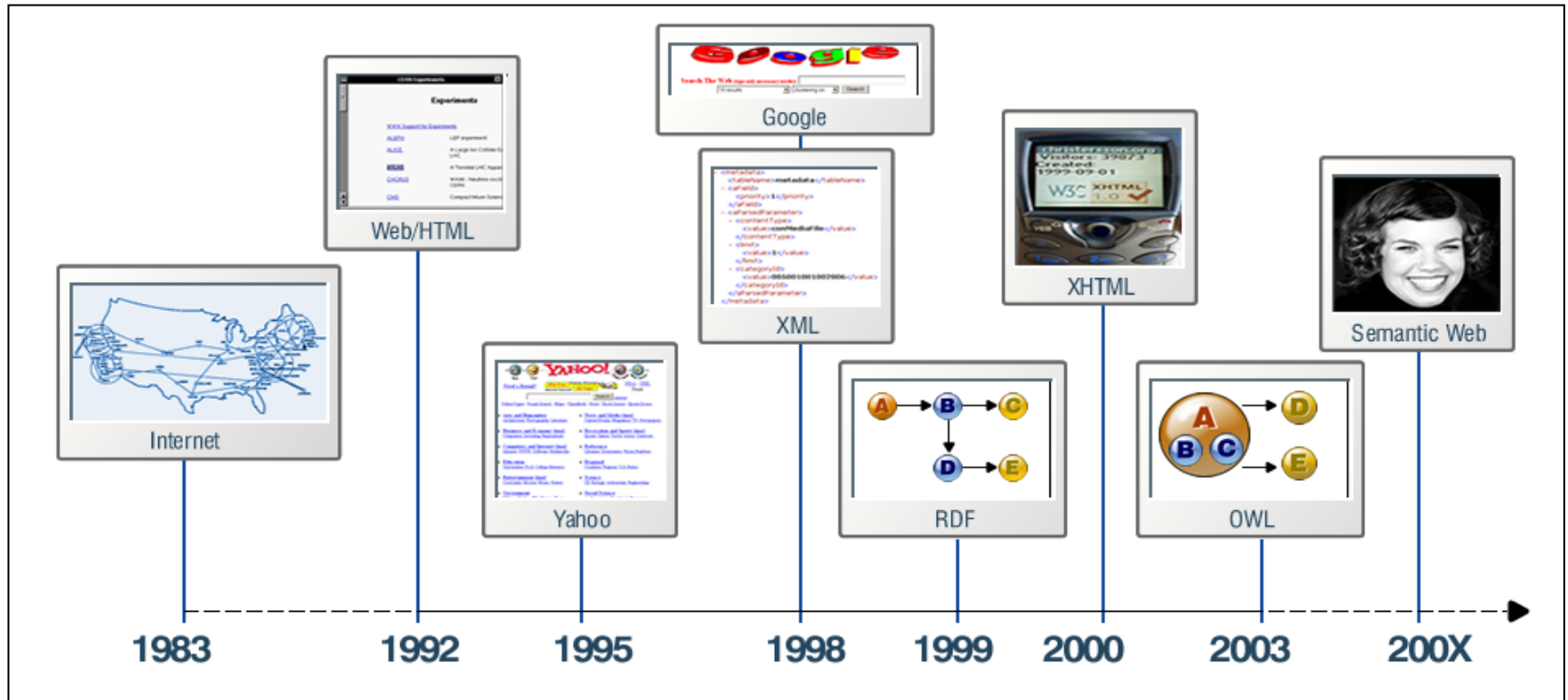
  - But needs to be converted to a new format.

It might be used for:

- Finding information/services
- Metadata
- An automation infrastructure
- Annotation
- Data mining
- …

Your guess is as good as mine.

Semantic Web: Past

Internet — 1983
Web/HTML — 1992
Yahoo — 1995
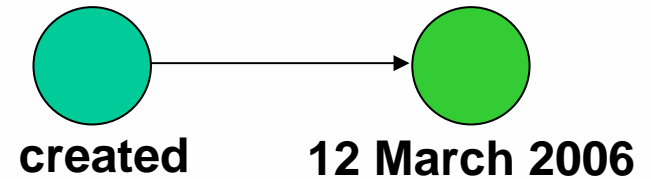Google — 1998
XML
RDF — 1999
XHTML — 2000
OWL — 2003
Semantic Web — 200X

Semantic Web: Past

- Metadata (attribute/value pair)

    - author   =   Jane Bloggs
    - created =   12 March 2006
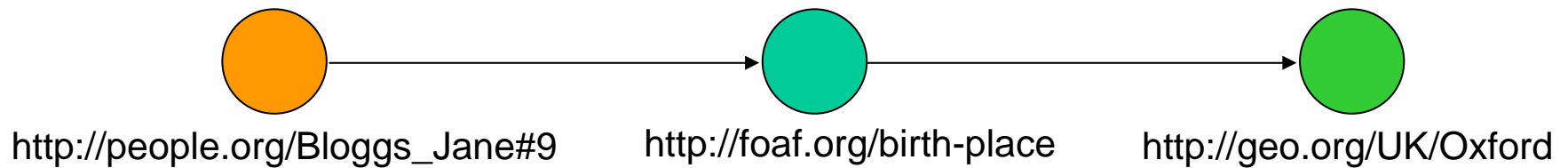
created   **12 March 2006**

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

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

Jane Bloggs   **born in**   **Oxford**

BOX | be creative. be innovative. be bold.

Semantic Web: Past

http://people.org/Bloggs_Jane#9   http://foaf.org/birth-place   http://geo.org/UK/Oxford
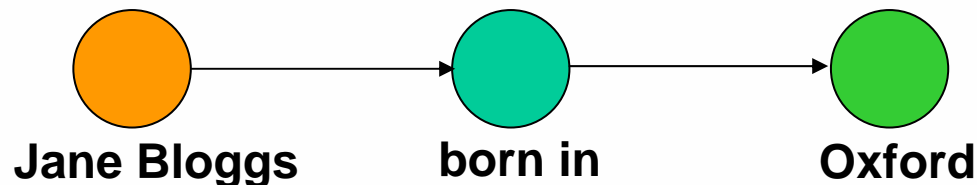
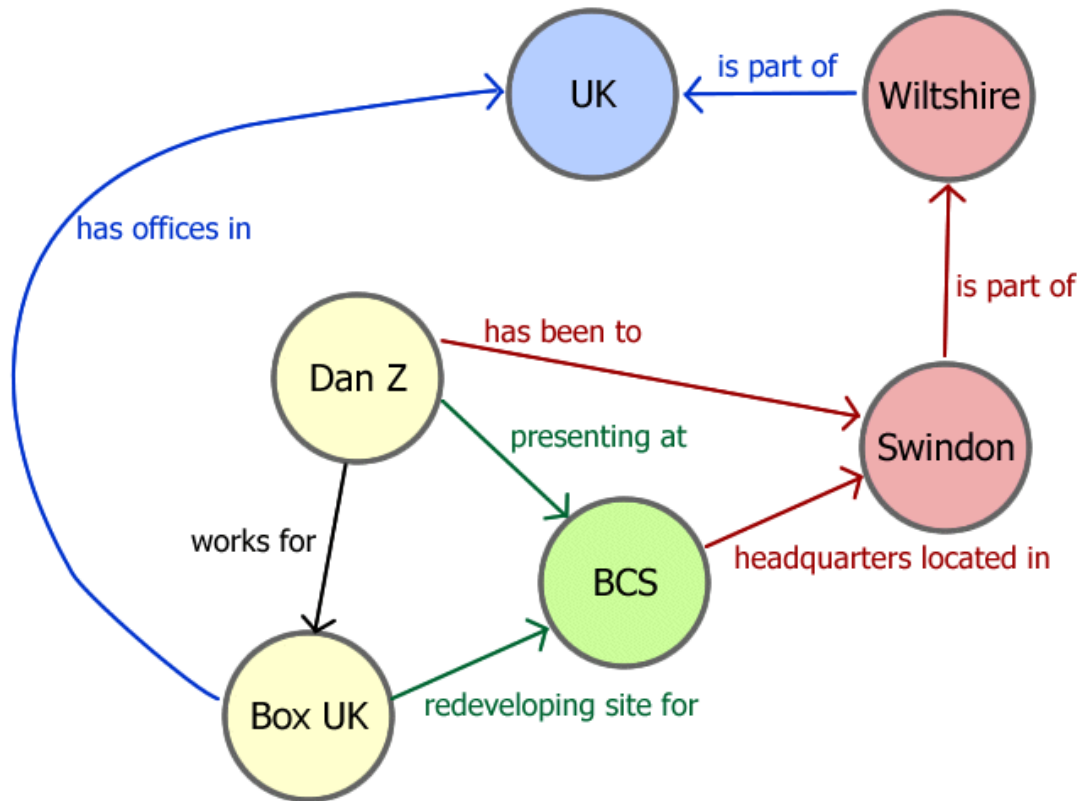URIs provide standard terms, RDF/XML a standard format.

```
<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>
```
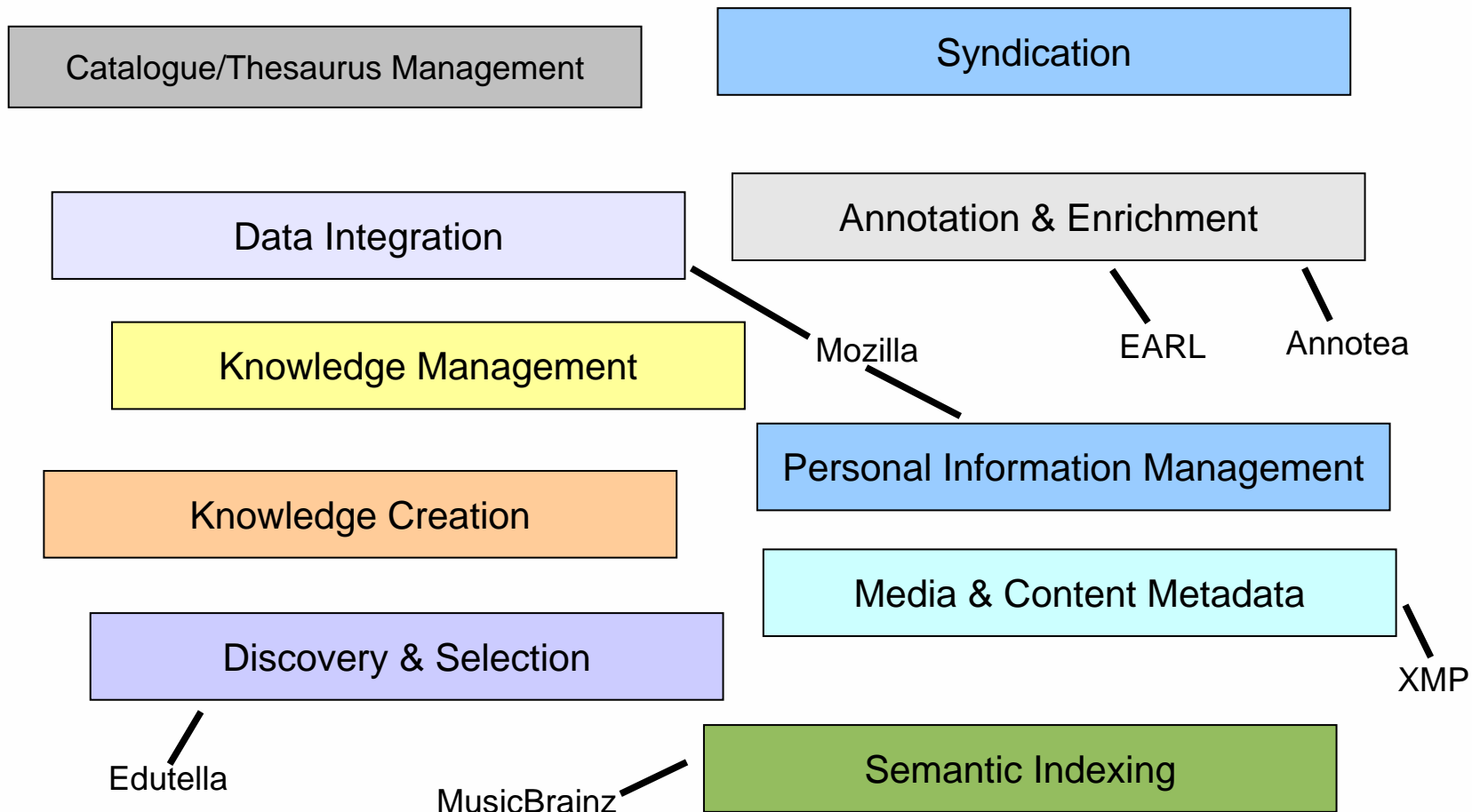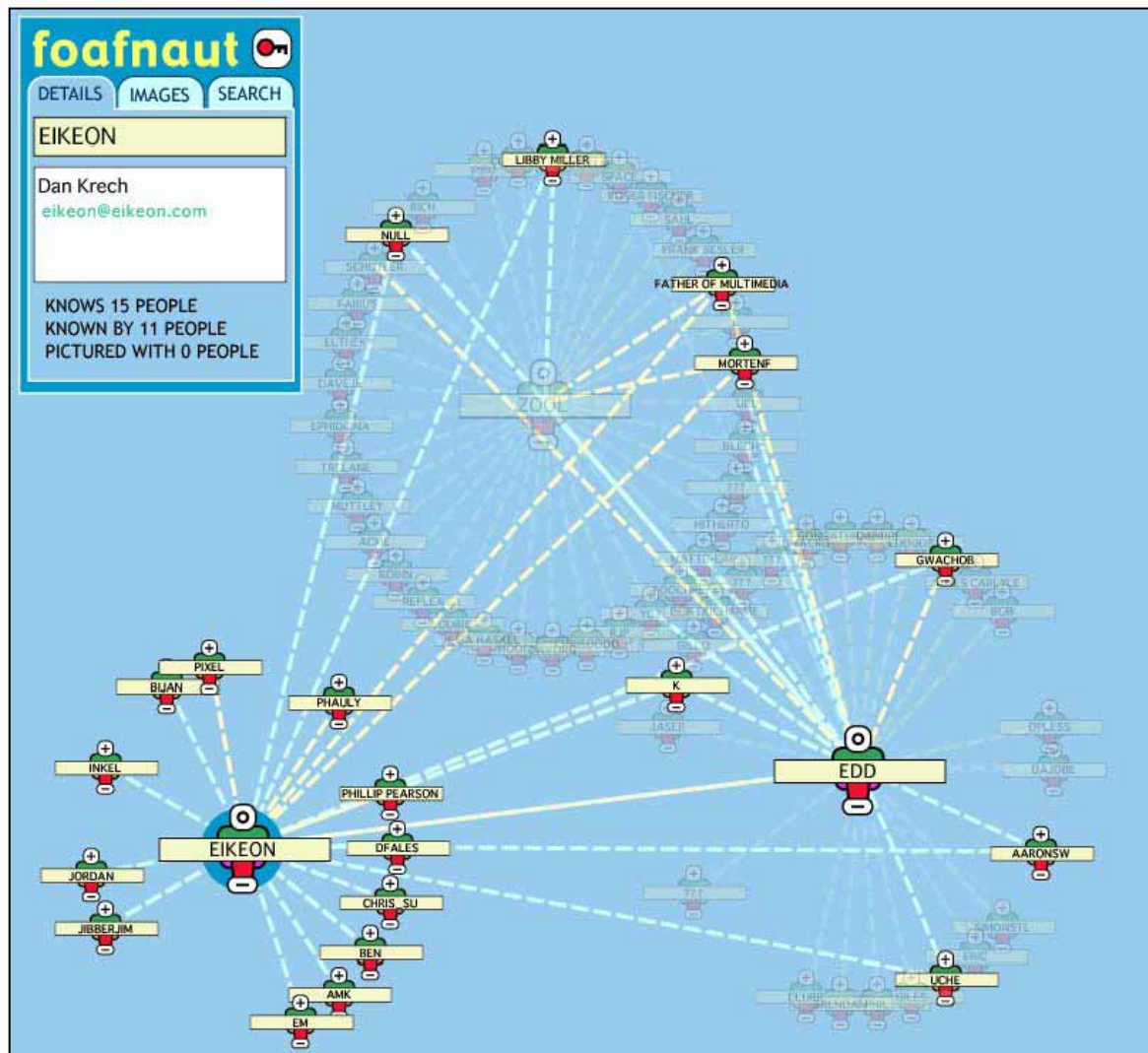
Semantic Web: Past

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:

Catalogue/Thesaurus Management

Syndication

Data Integration

Annotation & Enrichment

Mozilla

EARL

Annotea

Knowledge Management

Personal Information Management

Knowledge Creation

Media & Content Metadata

Discovery & Selection

XMP

Edutella

MusicBrainz

Semantic Indexing

Semantic Web: Present

Semantic Web: Present

Semantic Web: Present

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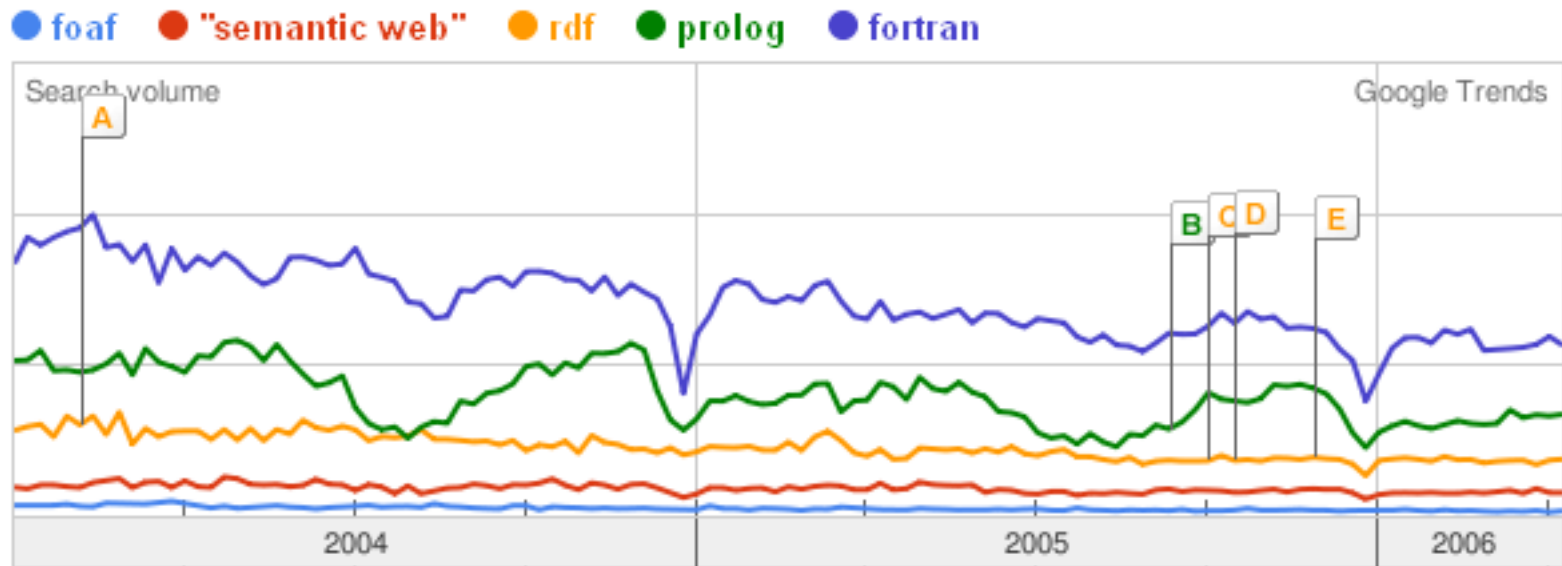
Interest in the Semantic Web flat-lined a long time ago.

Semantic Web: Present
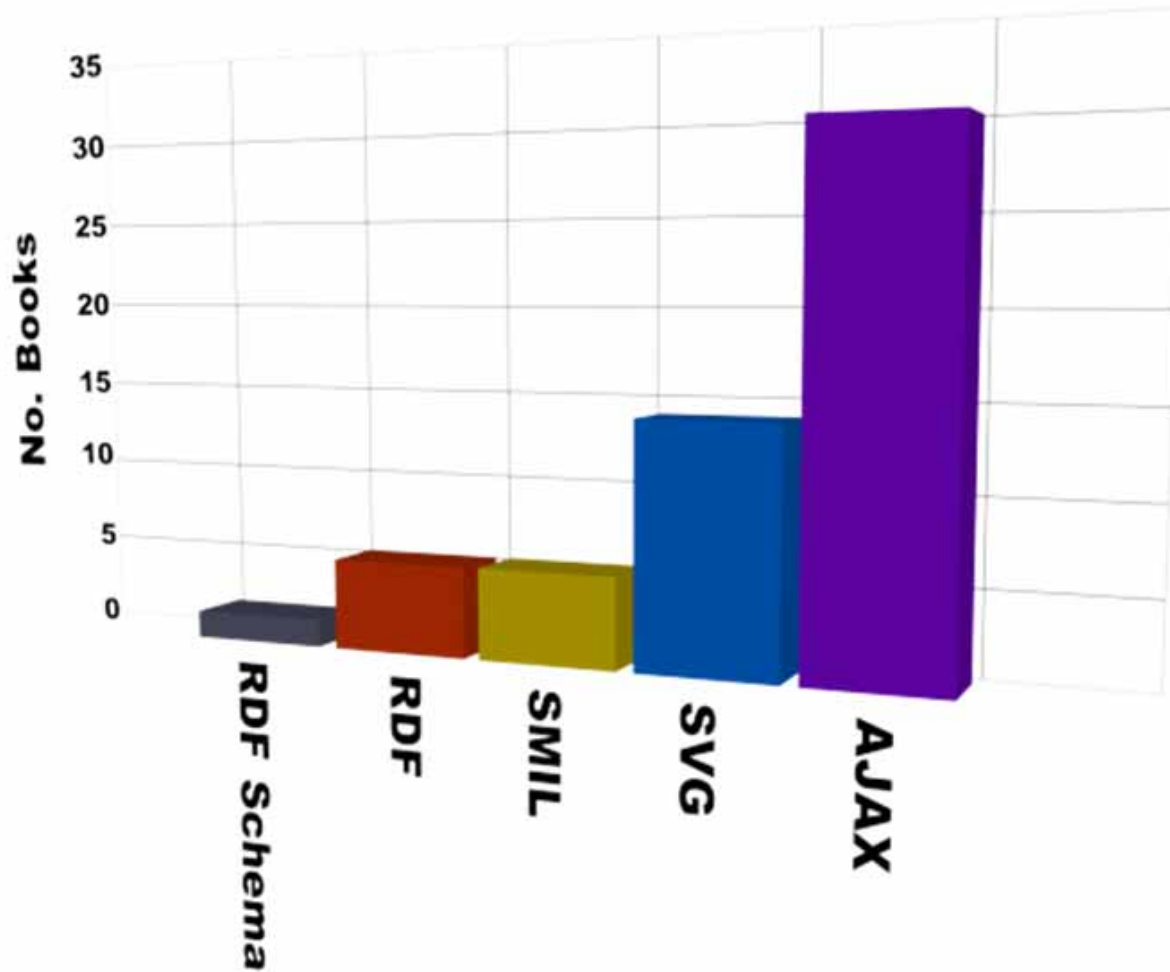
Books in Amazon.com under *Computers & Internet* category.

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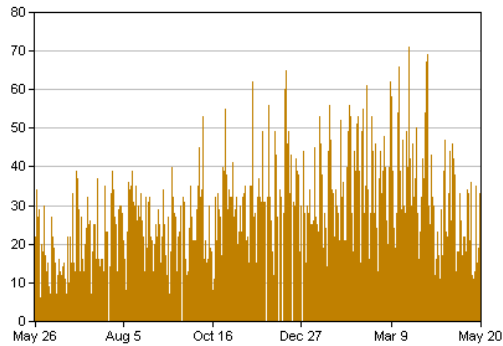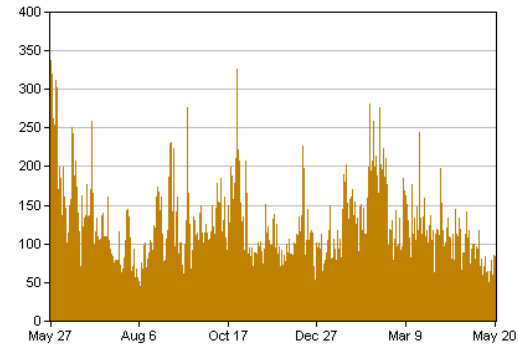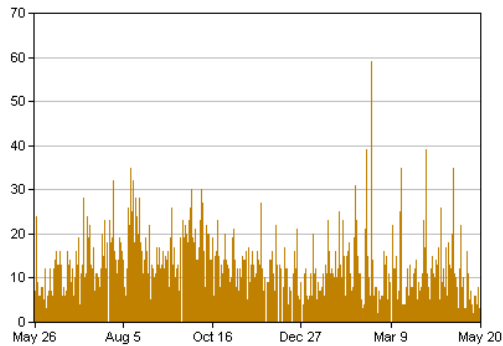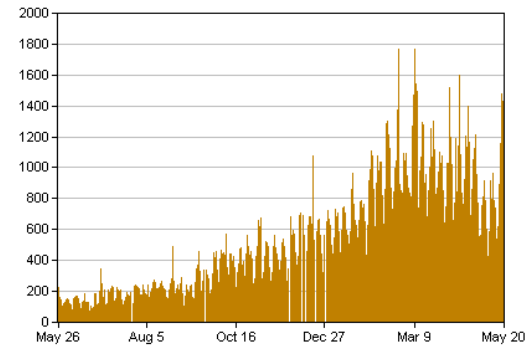
Blogs by subject, per day, for last year (source: technorati.com)

The 7 (f)laws of the Semantic Web.

Not all Semantic Web data are created equal.

The 1ˢᵗ (f)law of the Semantic Web
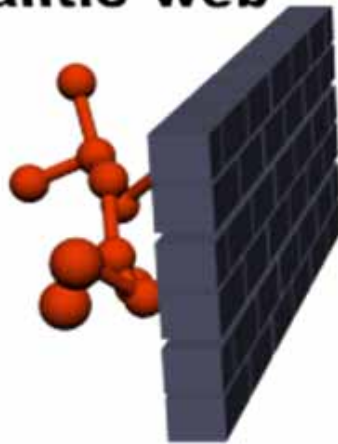
http://bbc.co.uk/news-12345   http://rss.org/subject   "Science/Nature"

BOX | be creative. be innovative. be bold.
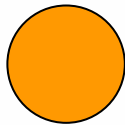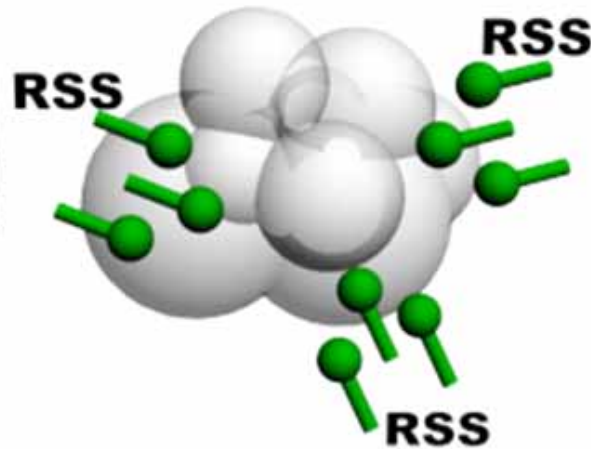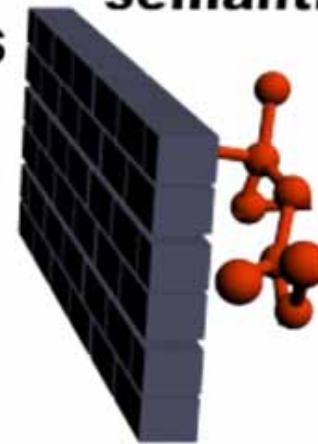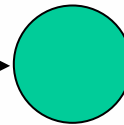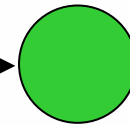
A technology is only as good as developers think it is.

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 evangalizing 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]

Complex Systems must be built from successively simpler systems.

BOX UK | be creative. be innovative. be bold.

The 3rd (f)law of the Semantic Web

'complexity'

'knowledge'

Web 2.0 Application

AJAX, DHTML

HTML, CSS, Javascript, XML

Semantic Web Application

Agents, Interference, Ontologies

RDF/XML, SPARQL, RDF Schema, URIs
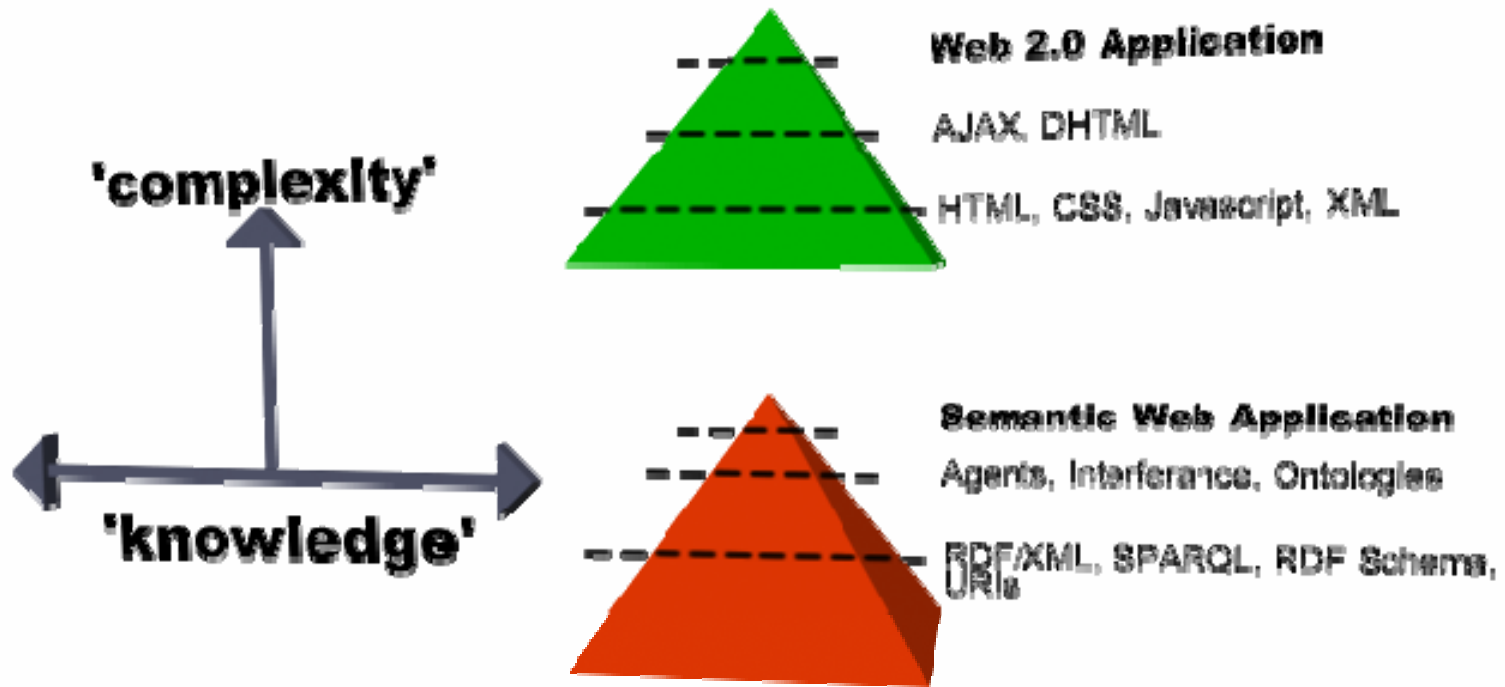
A new solution should stop an obvious pain.

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

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.

You don't need an **Ontology of Everything**.

But it would help.

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
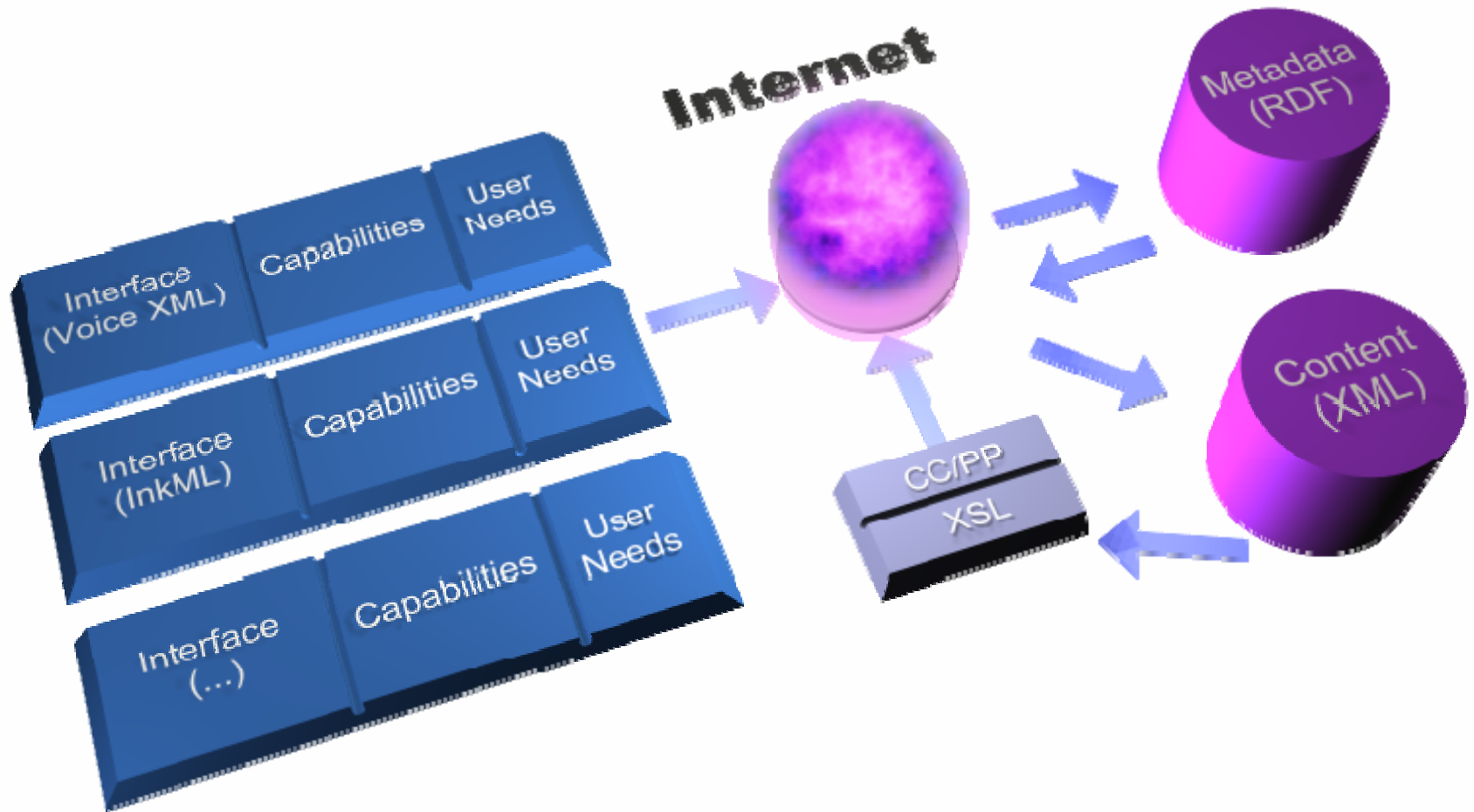
Semantic Web: Future

- 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!

BOX^{UK} | be creative. be innovative. be bold.

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/