OWL: Path to Massive Deployment

Dean Allemang

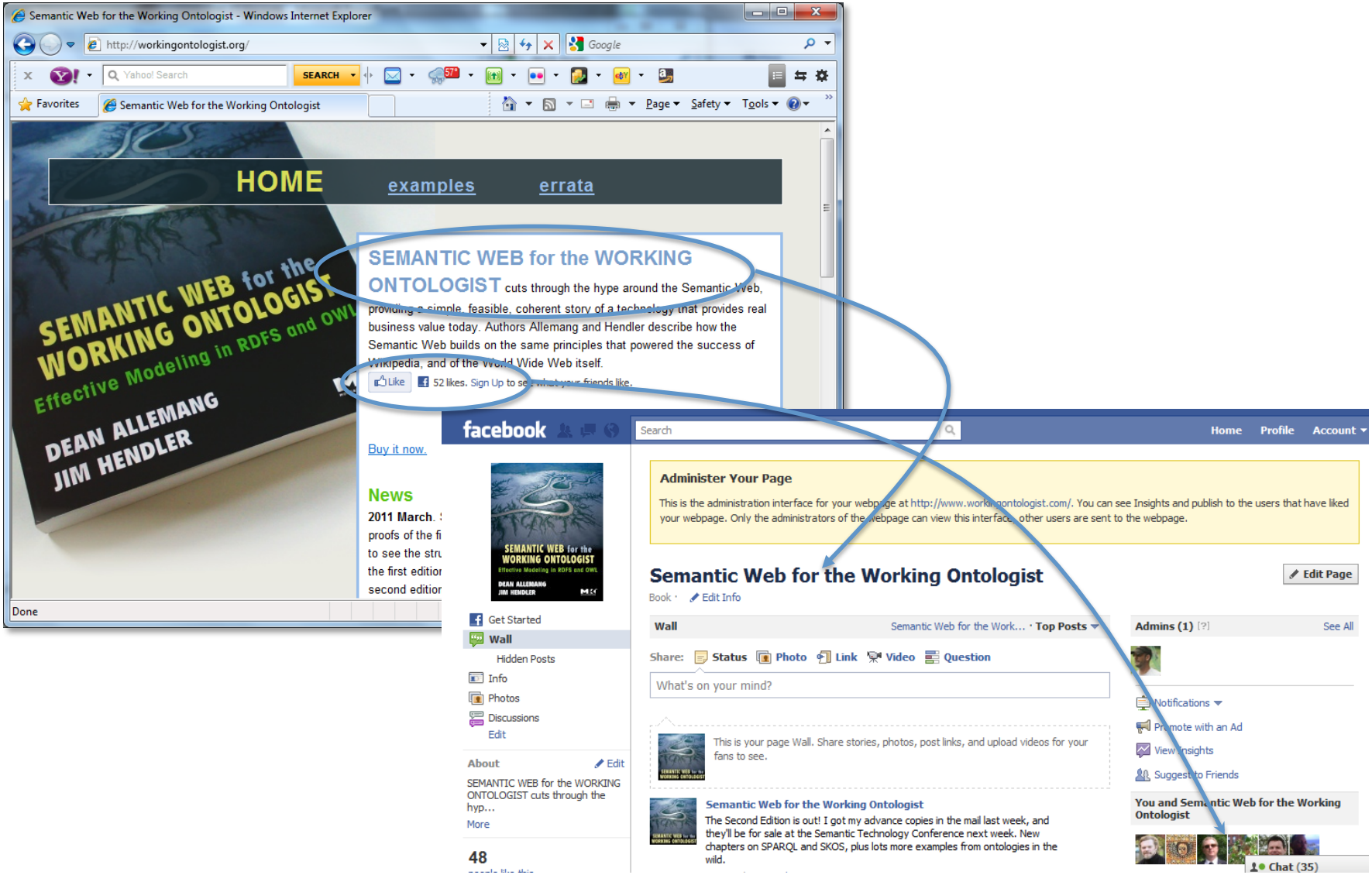
Chief Scientist, TopQuadrant Inc.

dallemang@topquadrant.com

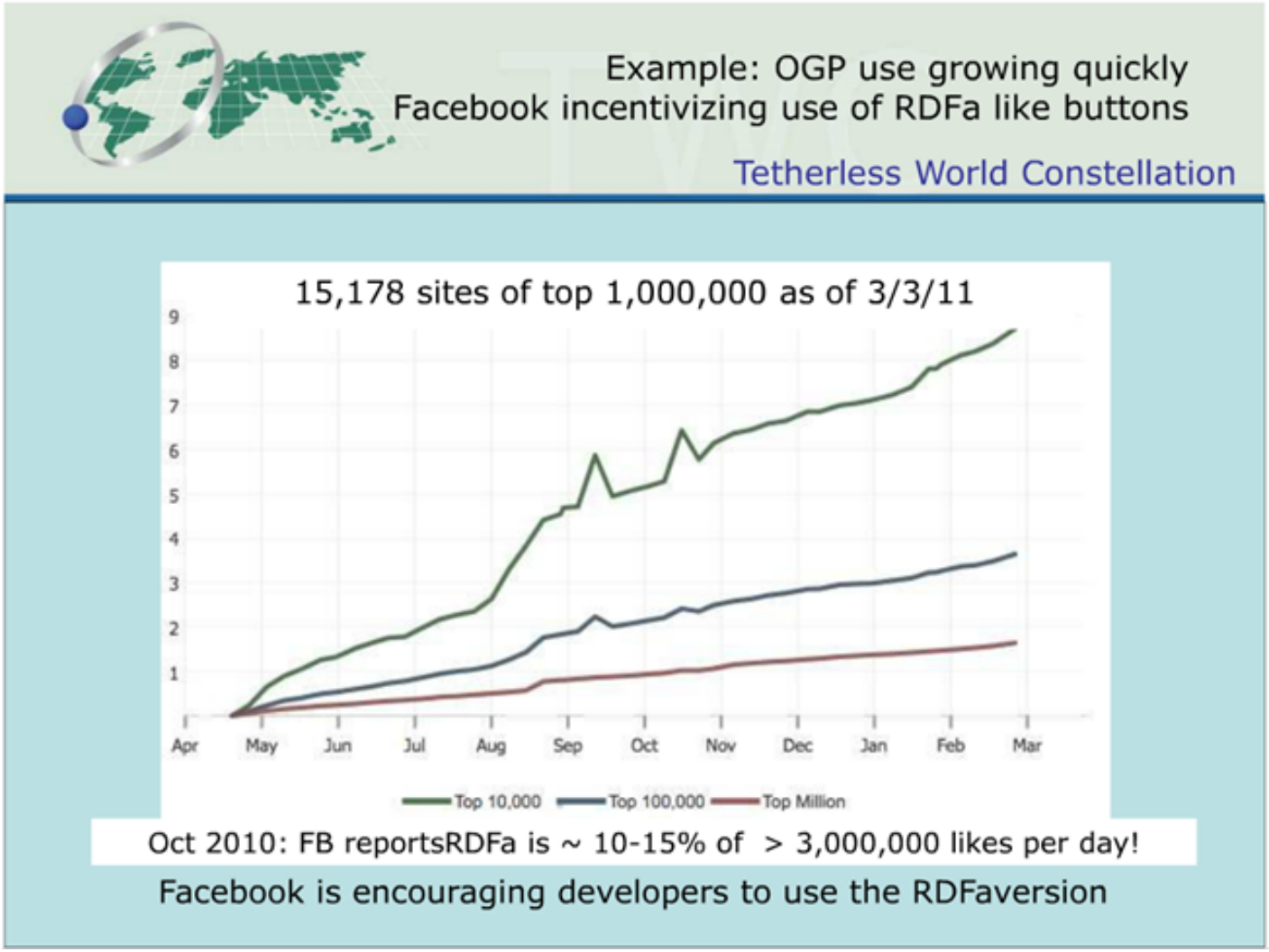
Web-Scale Deployment

- Number of pages
- Amount of Data
- Awareness
  - "I'm a Web Developer" "Have you heard of X?"
  - For X = Java, Ruby, HTML5, Semantic Web, RDF, SPARQL, Jena, OWL, Linked data, …

Facebook Open Graph Protocol



Facebook and OGP



Source: Jim Hendler

OGP and Simplicity

```
<meta property="og:latitude" content="37.416343"/>
<meta property="og:longitude" content="-122.153013"/>
<meta property="og:title" content="The Rock"/>
 <meta property="og:type" content="movie"/>
<meta property="og:url" content="http://www.imdb.com/title/
    tt0117500/"/>
<meta property="og:image" content="http://ia.media-imdb.com/
    rock.jpg"/>
 <meta property="og:site_name" content="IMDb"/>
<meta property="fb:admins" content="USER_ID"/>
<meta property="og:description" content="A group of U.S. Marines,
    under command of a renegade general, take over Alcatraz and
    threaten San Francisco Bay with biological weapons."/>
```

geo:lat, geo:long?

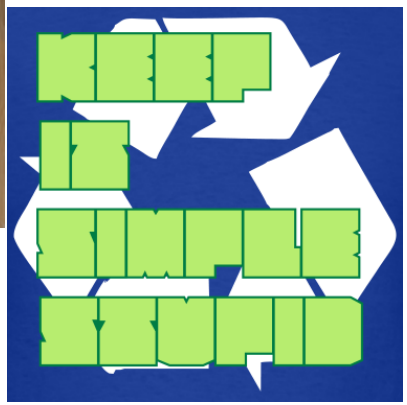
dc:title ?

foaf:depiction ?

rdfs:description ?
skos:definition ?

KISS

KISS
keep it simple...



Good Relations

- Vocabulary for eCommerce
- Lets vendors describe their products, services, prices, etc.
- Makes this data mergeable, dynamic, queryable, etc.

Source: Martin Hepp at http://purl.org/goodrelations

OWL Success Stories: Good Relations

site	pages	triples
overstock.com	1,000,000	100,000,000
CSNstores*	2,000,000	20,000,000
BestBuy*	500,000	5,000,000
Amazon*	20,000,000	4,000,000,000
Eurobau*		60,000,000
O'Reilly*	25,000	2,000,000
Bitmunk*		11,000,000

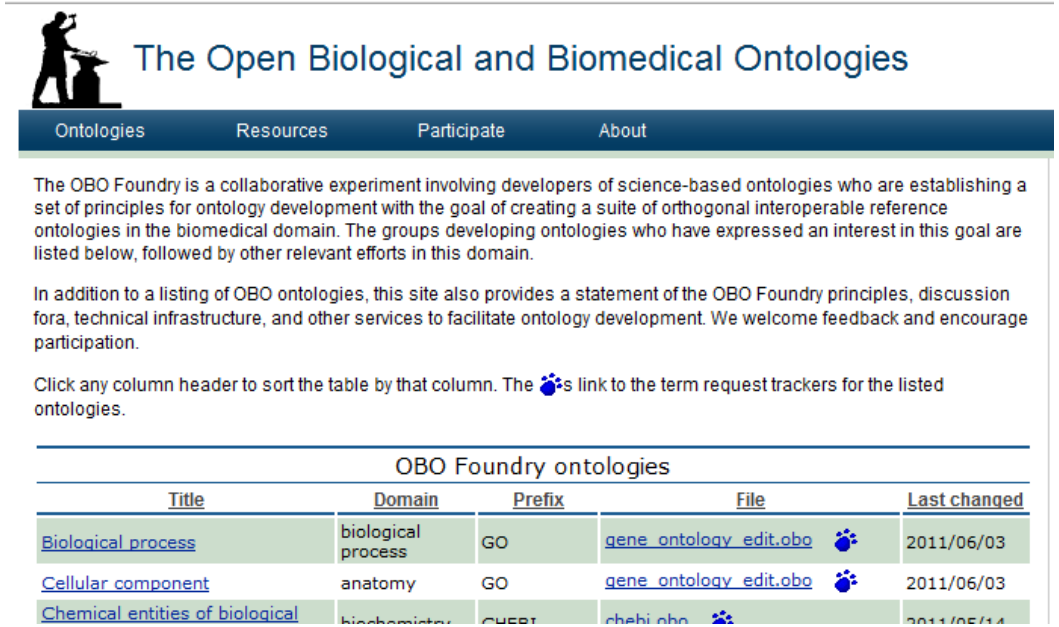
*Source: Good Relations wiki

SKOS (Simple Knowledge Organization System)

- System for managing controlled vocabularies
- 28 Vocabularies on the W3C page, including:
  - Dewey decimal system
  - Library of Congress
  - United Nations Agrovoc
  - Many more not on this page . . .
- Each vocabulary is referenced from sometimes thousands of sites

Open Biological and Biomedical Ontologies (OBO)

- Over a hundred curated models
- Some are quite large
  - CHEBI > 1 million triples
  - Gene Ontology ~ 2.5 million triples
- Over 100 million triples – that's getting to scale!



The Open Biological and Biomedical Ontologies

Ontologies | Resources | Participate | About

The OBO Foundry is a collaborative experiment involving developers of science-based ontologies who are establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain. The groups developing ontologies who have expressed an interest in this goal are listed below, followed by other relevant efforts in this domain.

In addition to a listing of OBO ontologies, this site also provides a statement of the OBO Foundry principles, discussion fora, technical infrastructure, and other services to facilitate ontology development. We welcome feedback and encourage participation.

Click any column header to sort the table by that column. The s link to the term request trackers for the listed ontologies.

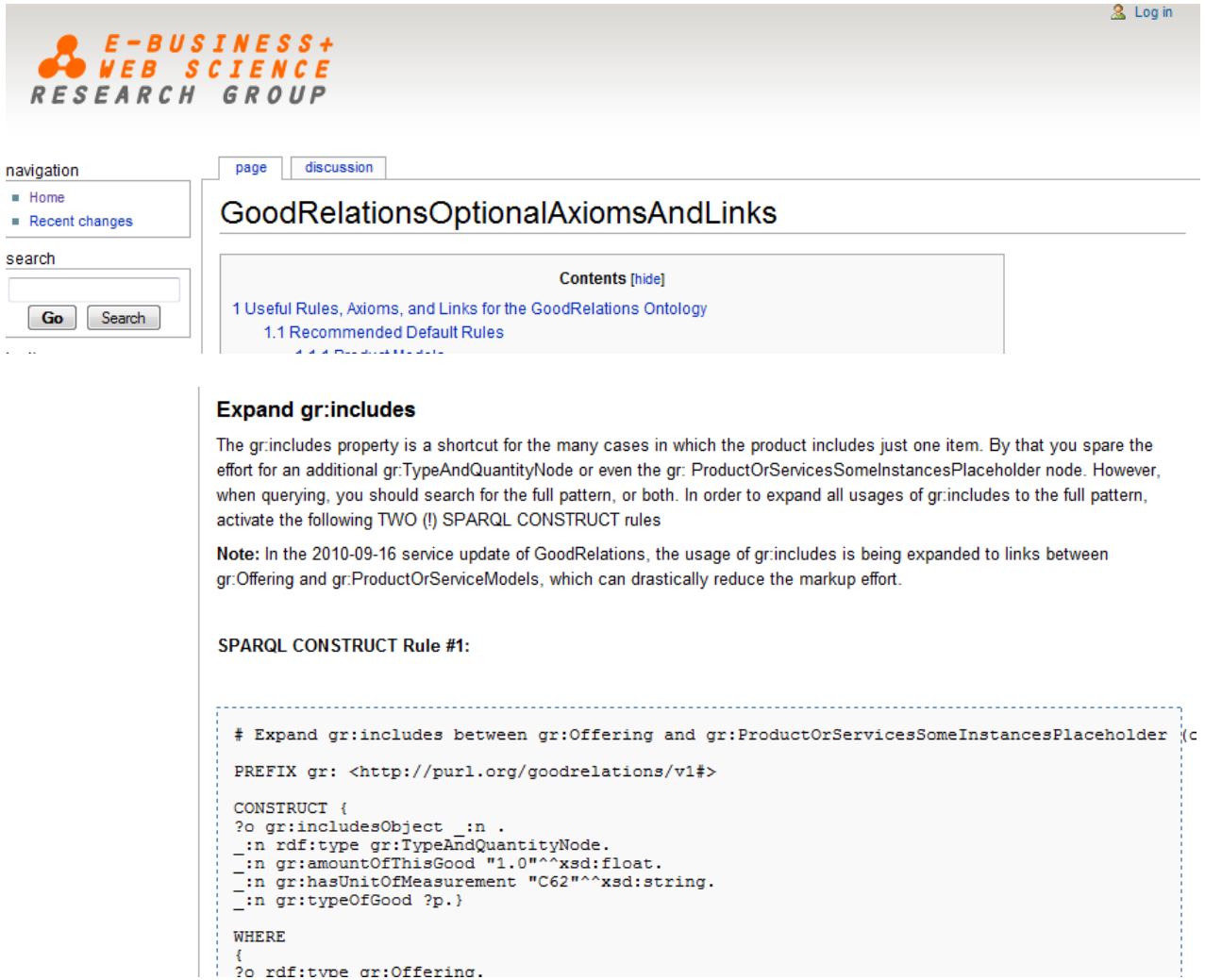
OBO Foundry ontologies

Title	Domain	Prefix	File	Last changed
Biological process	biological process	GO	gene_ontology_edit.obo	2011/06/03
Cellular component	anatomy	GO	gene_ontology_edit.obo	2011/06/03
Chemical entities of biological	biochemistry	CHEBI	chebi.obo	2011/05/14

OWL Utilization in common resources

	Good Relations	SKOS	OBO (CHEBI)
AnnotationProperty	X	X	X
Class	X	X	X
DatatypeProperty	X	X	
DeprecatedProperty	X		
FunctionalProperty		X	
ObjectProperty	X	X	X
Ontology	X	X	X
Restriction			X
SymmetricProperty	X	X	
TransitiveProperty	X	X	X
disjointWith	X	X	
inverseOf	X	X	
onProperty			X
someValuesFrom			X
unionOf	X	X	
versionInfo	X		

Optional Axioms in Good Relations – in SPARQL



Log in

E-BUSINESS+ WEB SCIENCE RESEARCH GROUP

navigation
- Home
- Recent changes

search
Go Search

page | discussion

GoodRelationsOptionalAxiomsAndLinks

Contents [hide]

1 Useful Rules, Axioms, and Links for the GoodRelations Ontology
1.1 Recommended Default Rules

Expand gr:includes

The gr:includes property is a shortcut for the many cases in which the product includes just one item. By that you spare the effort for an additional gr:TypeAndQuantityNode or even the gr: ProductOrServicesSomeInstancesPlaceholder node. However, when querying, you should search for the full pattern, or both. In order to expand all usages of gr:includes to the full pattern, activate the following TWO (!) SPARQL CONSTRUCT rules

Note: In the 2010-09-16 service update of GoodRelations, the usage of gr:includes is being expanded to links between gr:Offering and gr:ProductOrServiceModels, which can drastically reduce the markup effort.

SPARQL CONSTRUCT Rule #1:

```
# Expand gr:includes between gr:Offering and gr:ProductOrServicesSomeInstancesPlaceholder (c

PREFIX gr: <http://purl.org/goodrelations/v1#>

CONSTRUCT {
?o gr:includesObject _:n .
_:n rdf:type gr:TypeAndQuantityNode.
_:n gr:amountOfThisGood "1.0"^^xsd:float.
_:n gr:hasUnitOfMeasurement "C62"^^xsd:string.
_:n gr:typeOfGood ?p.}

WHERE
{
?o rdf:type gr:Offering.
```

Namespaces used in Overstock.com product pages

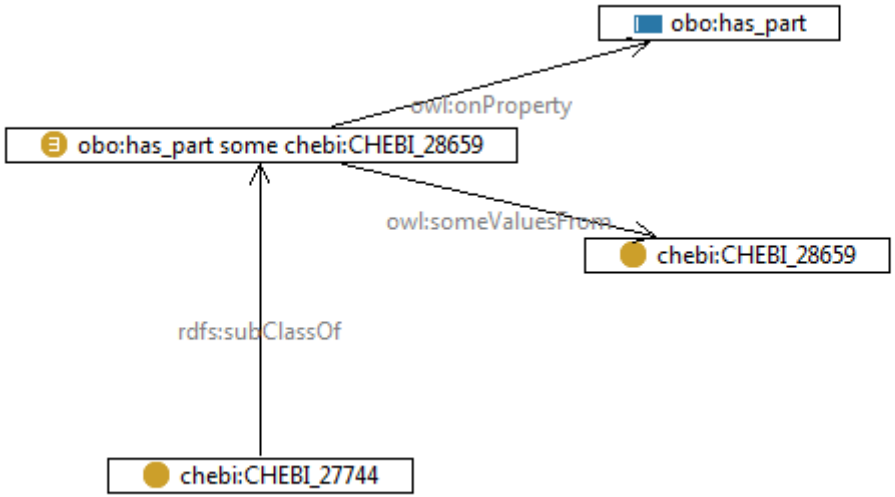
- The actual overstock.com product pages include commerce data about their products. This data uses:
  - OGP (image etc.)
  - GR (Good Relations)
  - FOAF (depiction)
  - RDFS (comment and label)
  - RDF review vocabulary (http://vocab.org/review/terms.html)
  - RDF data vocabulary (http://rdf.data-vocabulary.org/)
  - Facebook markup (http://www.facebook.com/2008/fbml#)

- OWL is missing – and that's a good thing!

Using Good Relations doesn't require learning OWL

How about OBO?

How do I say “glyphosate is made of phosphorus” in CHEBI?



“glyphosate is a subclass of the set of things that have some part that is phosphorus”

figure made with TopBraid Composer™

How to Make $$$ in OWL Today

- Make OWL look really hard ("Drinking from a firehose" "Very Complicated")
  - Sell a training course to help people out
  - Write a book about OWL
  - Sell OWL consulting ("don't try this at home!")

How to Make $$$ in OWL Tomorrow

OWL must not be mysterious

Who is interested in the Semantic Web?

- Database architect
- Java programmer
- Library scientist
- Data modeler
- IT Specialist
- Enterprise Architect
- Knowledge Manager
- Entrepreneur
- Vocabulary Specialist
- Product manager
- Systems Engineer
- Consultant
- Software Engineer
- Taxonomist
- Scientist (geographer, chemist, geologist, linguist)

What is hard about OWL?

- Open World Assumption
  - Data applications have all been closed world
  - Usually I want to use that data again anyway
- Domain and Range
- Inferencing vs. Processing
  - Declarative vs. procedural
  - Programmers want control
  - Data folks want queries
- Logic
  - allValuesFrom can mean “none”

“all” can mean “none”

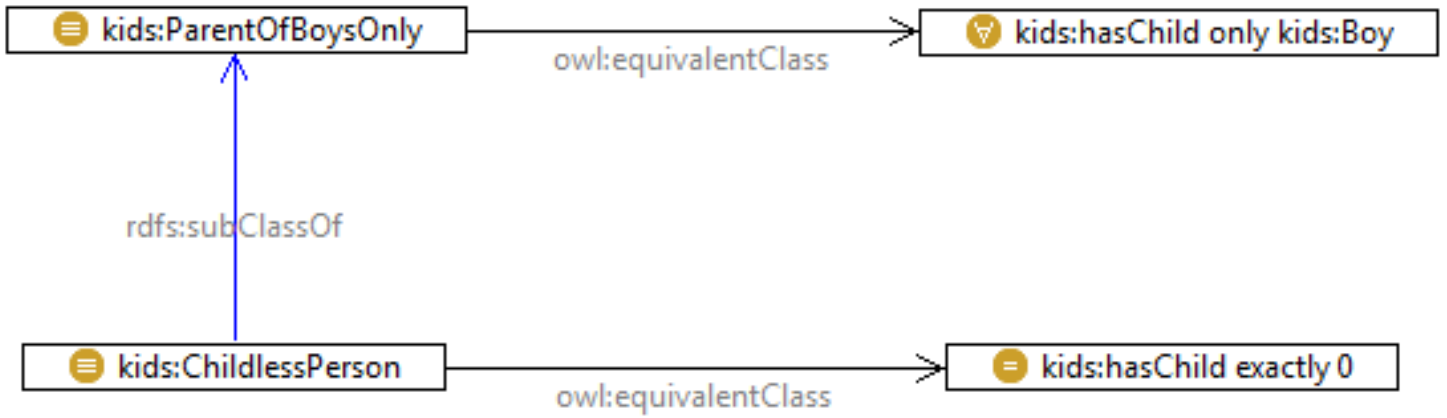
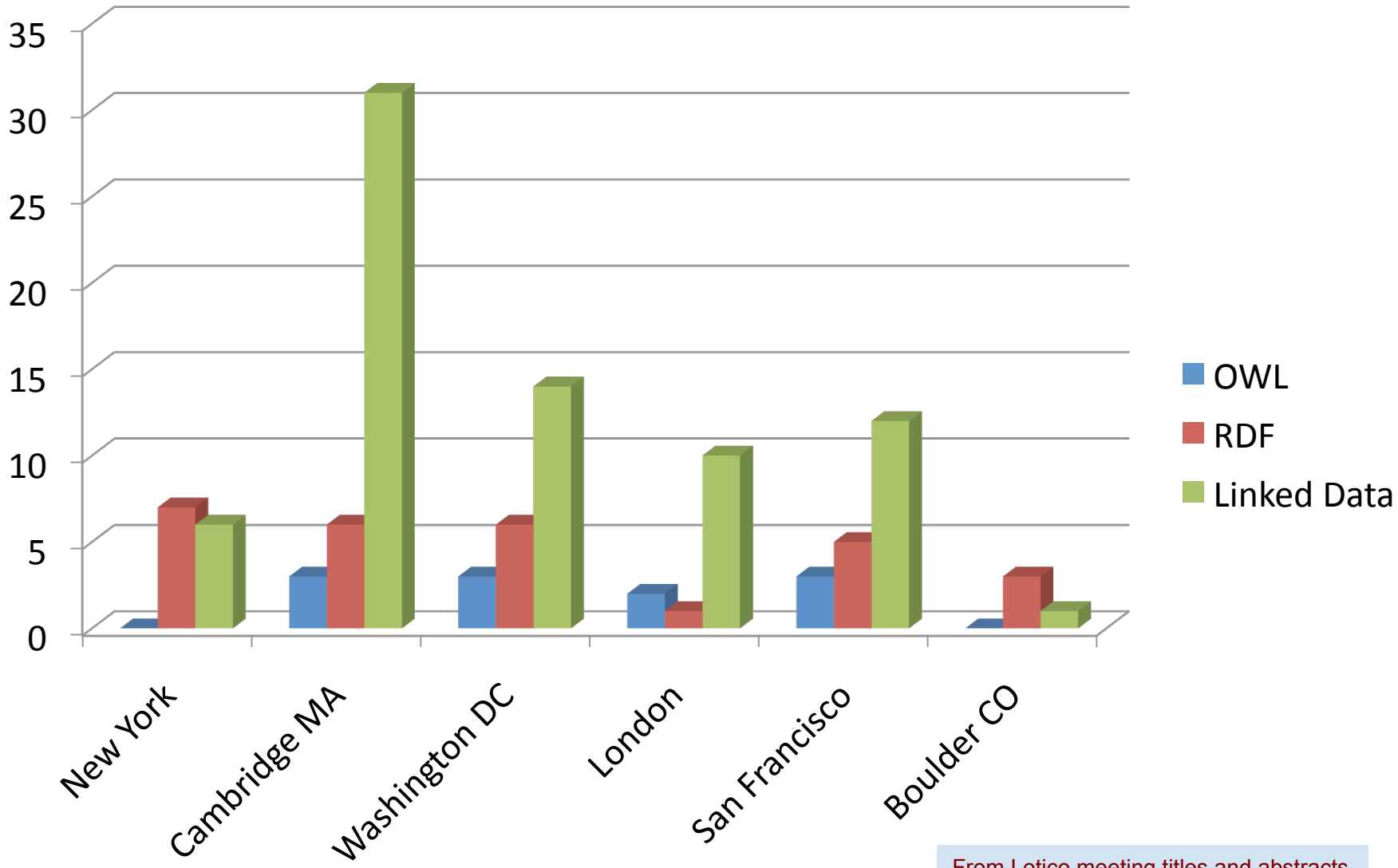


figure made with TopBraid Composer™

Lotico Topics of Interest



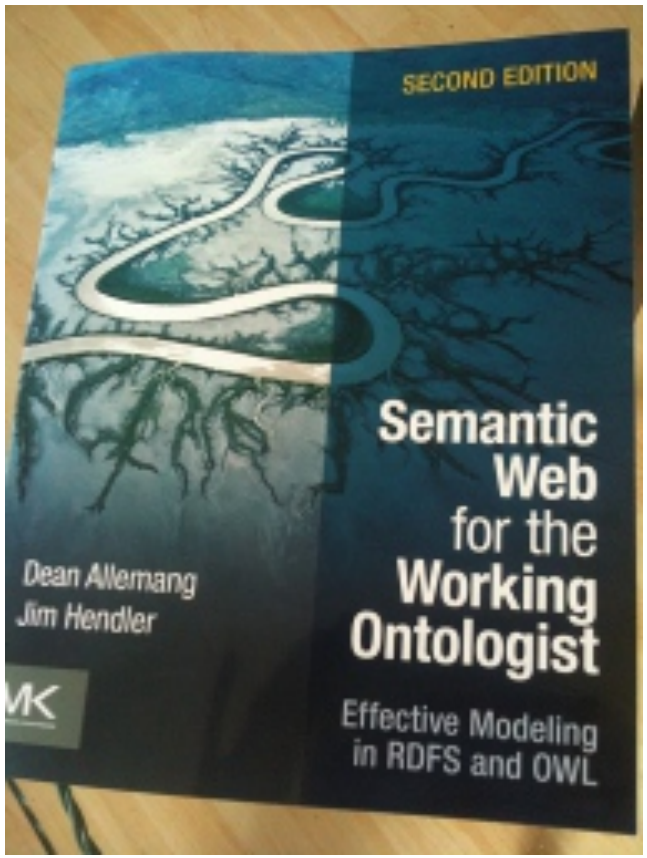
From Lotico meeting titles and abstracts

Demystification of OWL

- You may be surprised by what other people don't know – and whether they are motivated to learn it!

- One of the most successful OWL deployments is one where end deployers don't need to know OWL – It can be done!

Read more



- Semantic Web for the Working Ontologist by Dean Allemang and Jim Hendler
- Second Edition available now!
- Features
  - Good Relations
  - QUDT
  - Lots more!