What's new on the Web of Linked Data?

Richard Cyganiak (DERI)

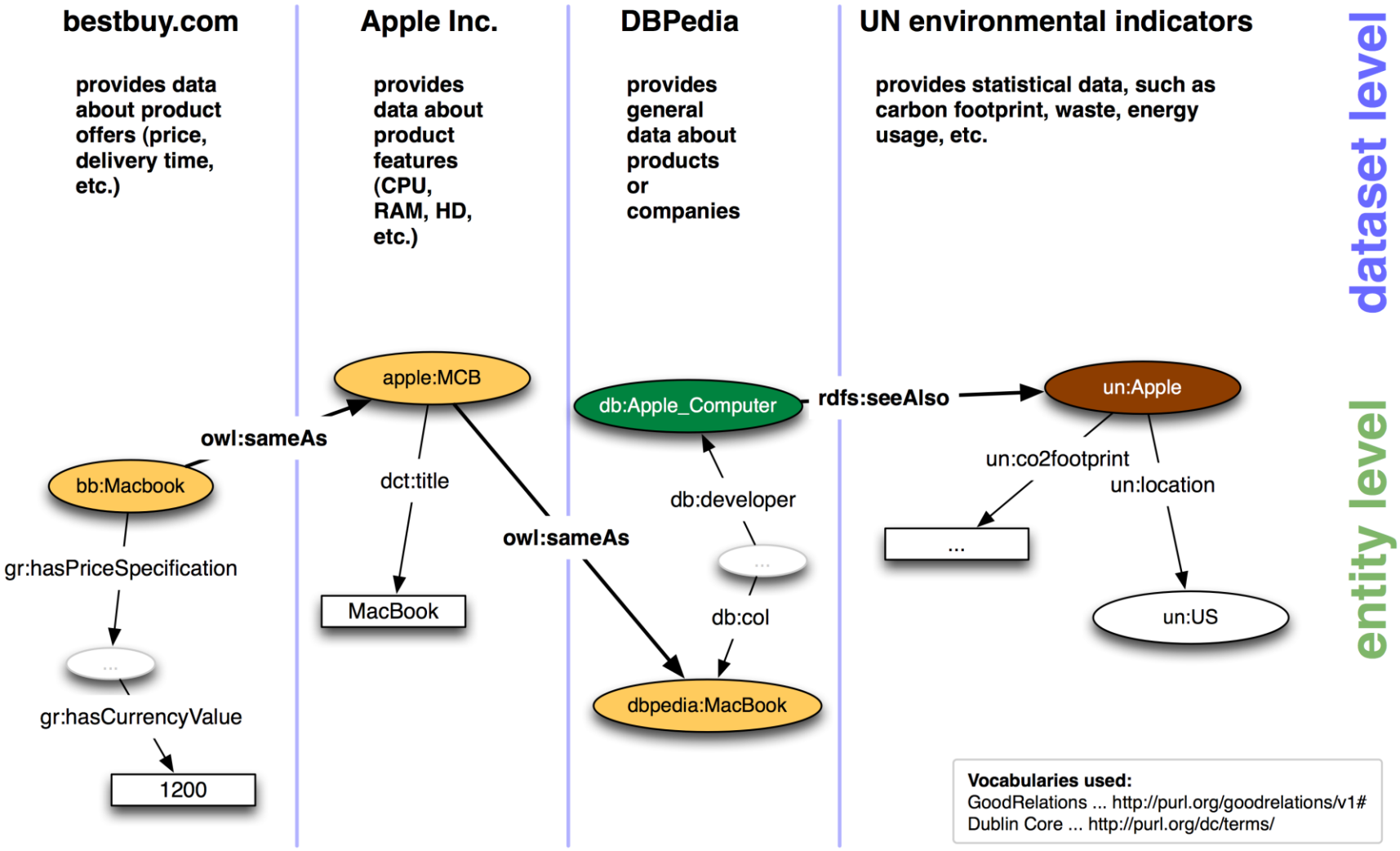
23 May 2013

Enabling **Networked** Knowledge

Agenda

- Short introduction to Linked Data
- Standard Vocabularies and Data Modelling for the Web of Data
- CSV as Linked Data: Linked CSV
- Provenance: W3C's PROV Model
- JSON as Linked Data: JSON-LD

Linked Data example



In one sentence

Linked data attempts to do for data what the World Wide Web did for documents

Linked data turns the abstract vision of the Semantic Web into concrete technical guidelines that can actually be implemented and deployed

In one sentence

Linked data is about surfacing the contents of databases on the Web

Linked data reduces the technical cost of agreeing on common representations for data, giving us a better shot at tackling the social part of that problem

How it works

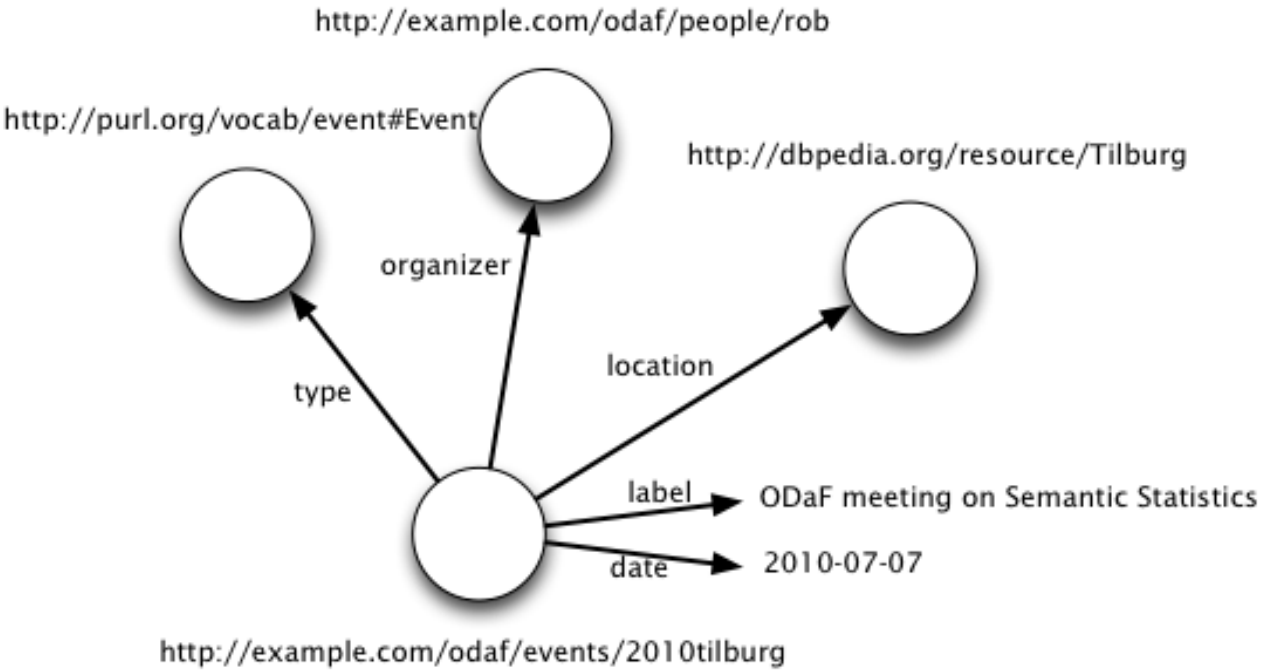
How it works

ODaF meeting on Semantic Statistics

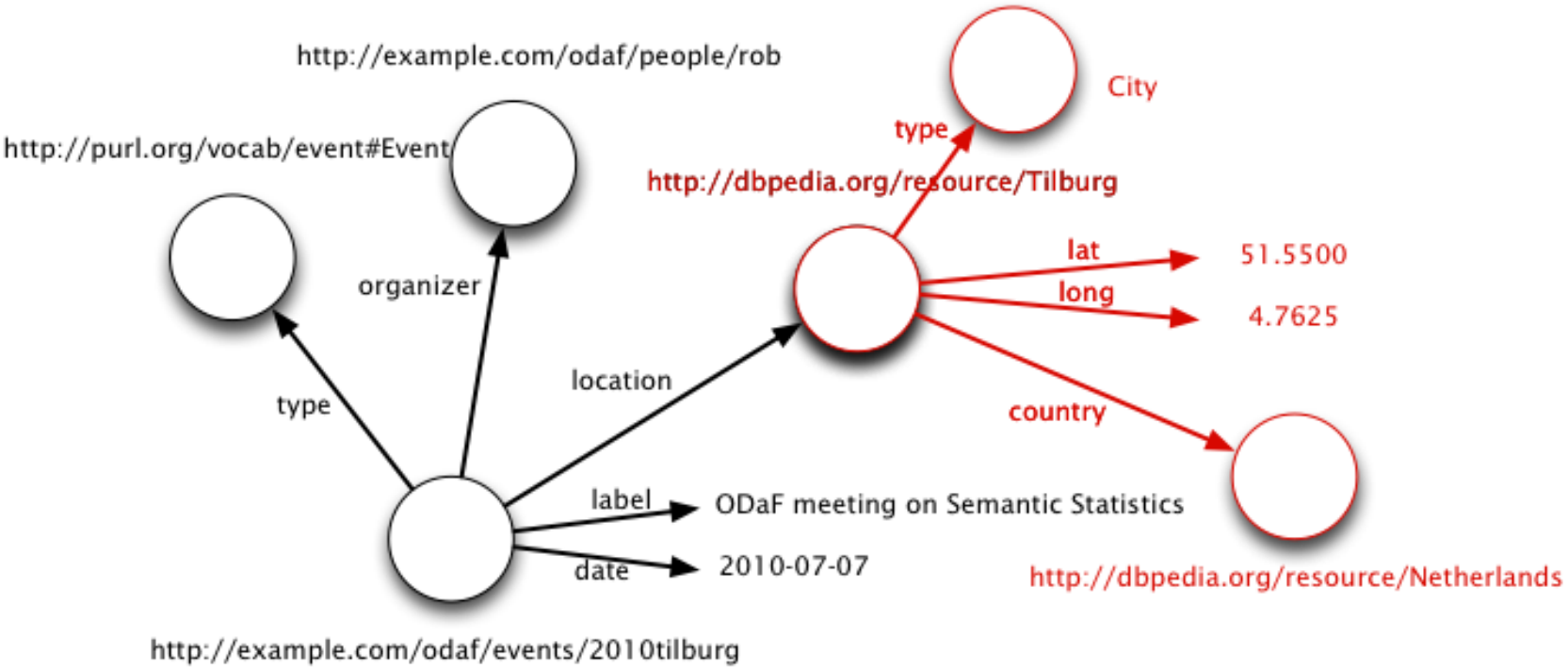
How it works

http://example.com/odaf/events/2010tilburg

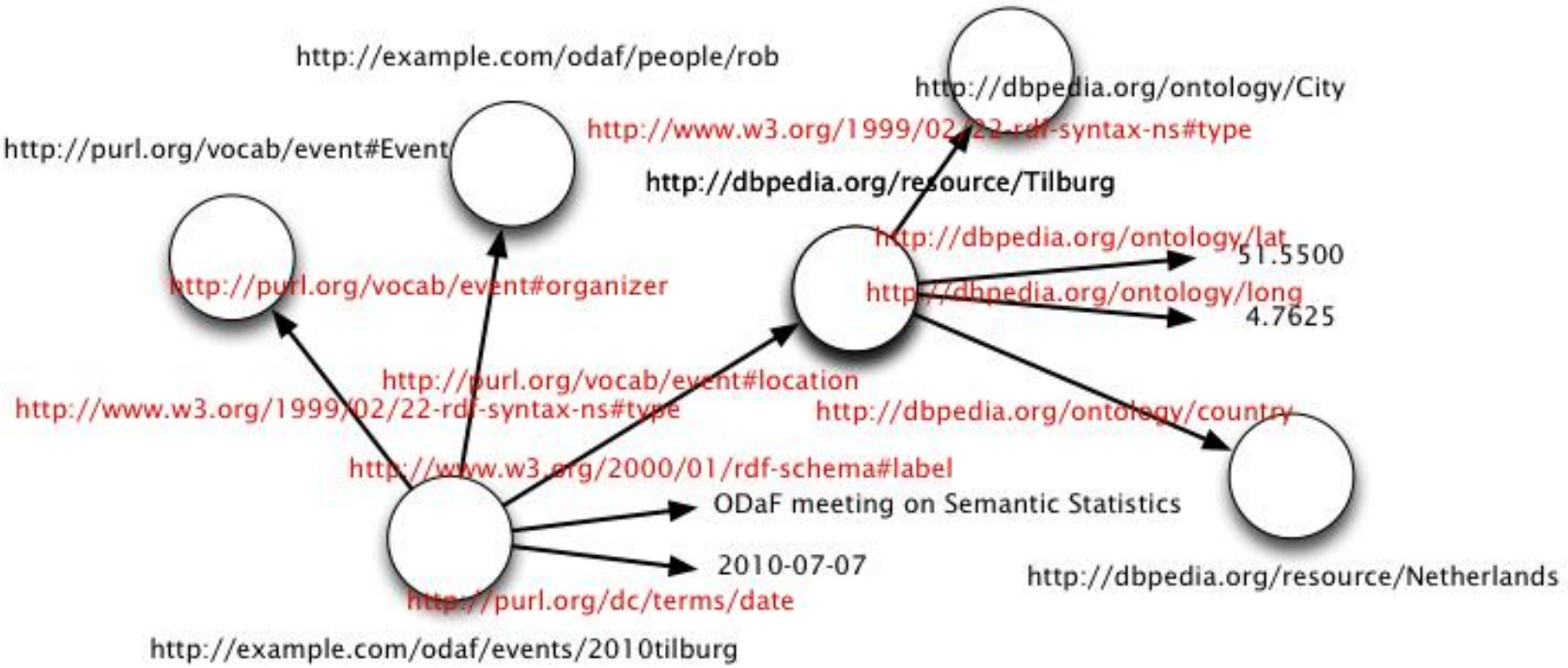
How it works


http://example.com/odaf/people/rob
http://purl.org/vocab/event#Event
http://dbpedia.org/resource/Tilburg
organizer
type
location
label
ODaF meeting on Semantic Statistics
date
2010-07-07
http://example.com/odaf/events/2010tilburg

How it works


http://example.com/odaf/people/rob
City
type
http://purl.org/vocab/event#Event
http://dbpedia.org/resource/Tilburg
lat
51.5500
long
4.7625
organizer
location
type
country
label
ODaF meeting on Semantic Statistics
date
2010-07-07
http://dbpedia.org/resource/Netherlands
http://example.com/odaf/events/2010tilburg

How it works


http://example.com/odaf/people/rob
http://dbpedia.org/ontology/City
http://www.w3.org/1999/02/22-rdf-syntax-ns#type
http://purl.org/vocab/event#Event
http://dbpedia.org/resource/Tilburg
http://dbpedia.org/ontology/lat
51.5500
http://purl.org/vocab/event#organizer
http://dbpedia.org/ontology/long
4.7625
http://purl.org/vocab/event#location
http://www.w3.org/1999/02/22-rdf-syntax-ns#type
http://dbpedia.org/ontology/country
http://www.w3.org/2000/01/rdf-schema#label
ODaF meeting on Semantic Statistics
2010-07-07
http://purl.org/dc/terms/date
http://dbpedia.org/resource/Netherlands
http://example.com/odaf/events/2010tilburg

URIs

Resolvable

Return HTML and RDF

Anyone with a web server can play

Giant Global Graph

Distributed publishing

Joined by different publishers using each others URIs

Enables crawling and discovery

Common data model

RDF—a graph consisting of triples

Standard tools for querying, storage etc

SPARQL query language

Shared vocabularies

Collections of classes and properties

Sometimes called “ontologies”

Vocabularies can be mixed and can co-exist

Popular examples: DC, FOAF, Schema.org

Re-usable URI sets

Publishers become “hubs”

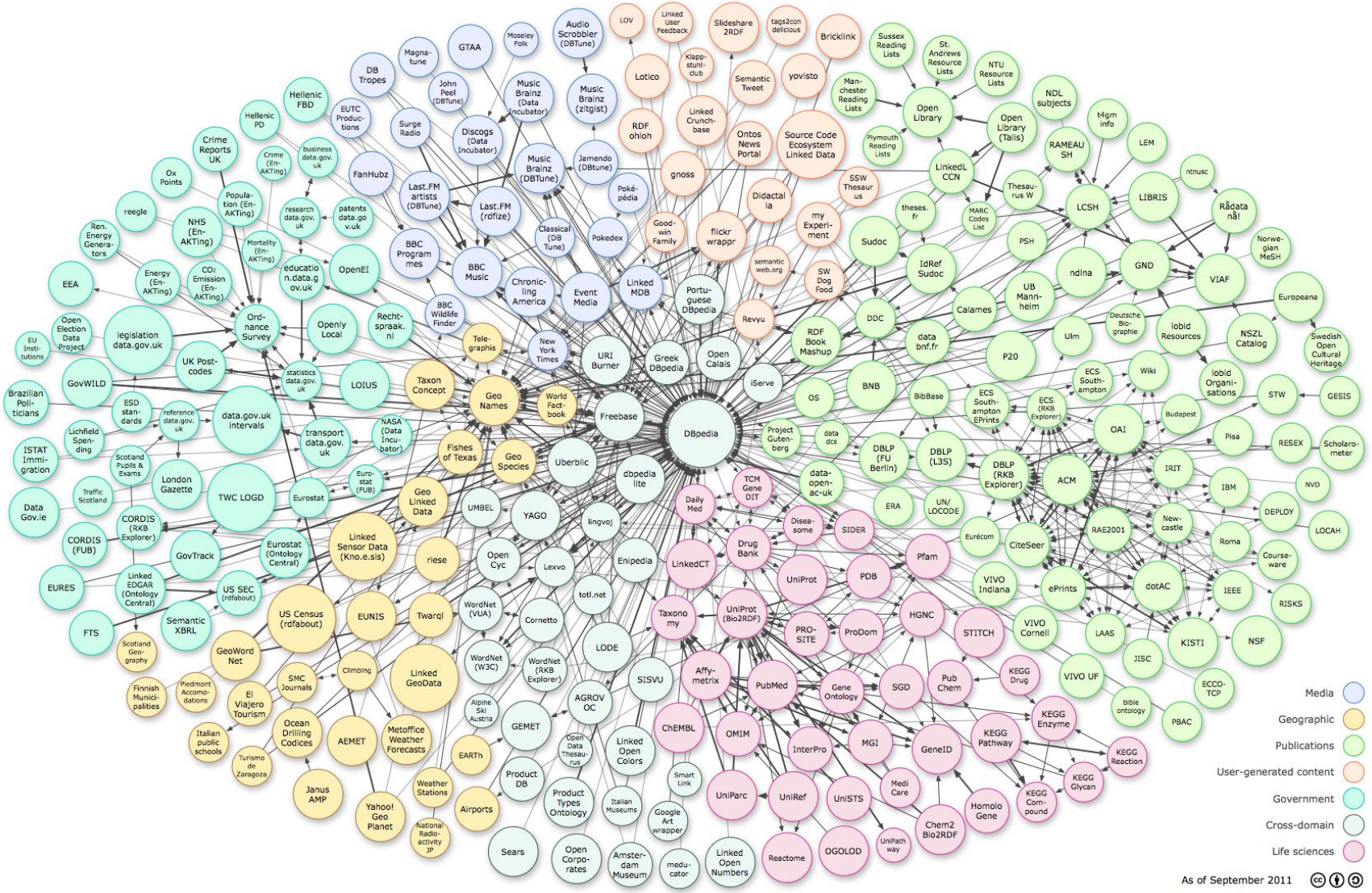
Greatly simplifies data integration

Examples: DBpedia, Geonames, UK government

Open Data, Read-Only

Because it's simpler

People are working on access control and read/write


Media
Geographic
Publications
User-generated content
Government
Cross-domain
Life sciences
As of September 2011

TimBL's 5-star plan

★ Make your data available on the Web

★★ Make it available as structured data
(e.g., an Excel sheet instead of image scan of a table)

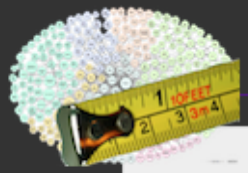
★★★ Use a non-proprietary format
(e.g., a CSV file instead of an Excel sheet)

★★★★ Use linked data format
(i.e., URIs to identify things, and RDF to represent data)

★★★★★ Link your data to other people's data to provide context
(i.e., re-use standard vocabularies and add cross-dataset links

Standard Vocabularies

- Live statistics on usage: stats.lod2.eu
- Bottom-up ontologies: Schema.org
- Bird's eye view on connections + reuse: lov.okfn.org



LODStats

Creating Knowledge out of Interlinked Data

LOD2

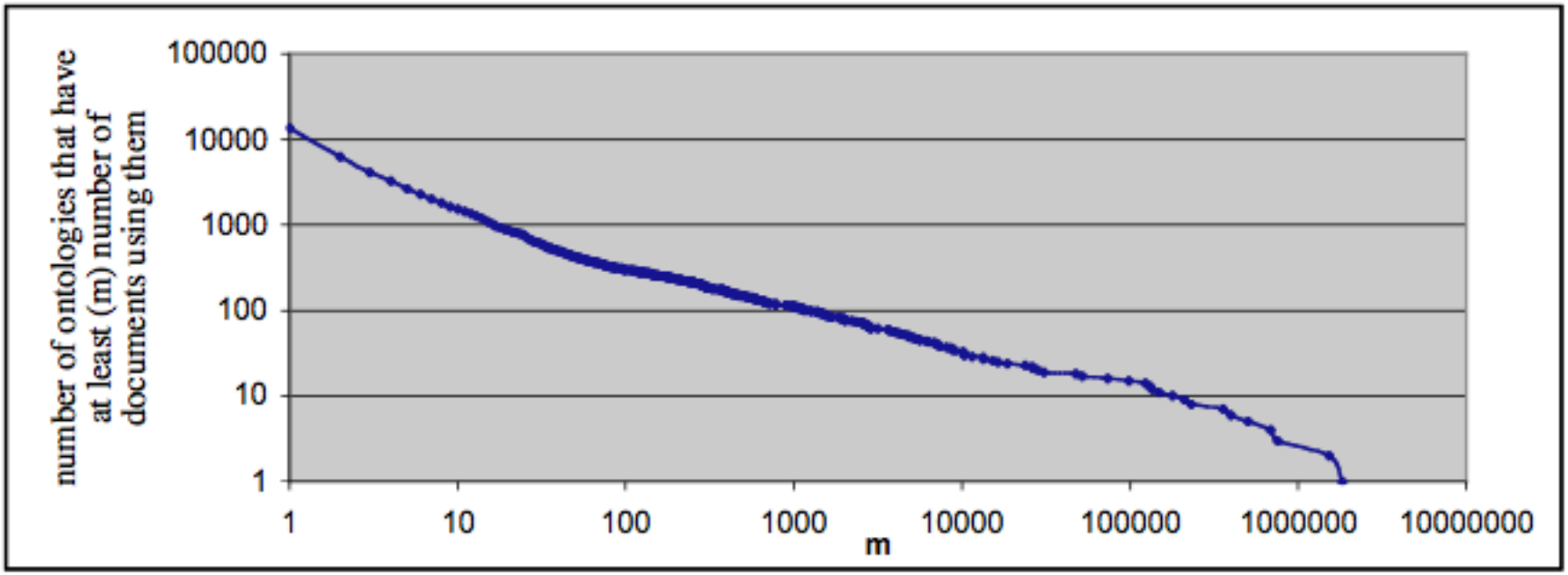
AKSW

499 vocabularies

Enter query here | Search

URI	Overall	Datasets
http://www.w3.org/1999/02/22-rdf-syntax-ns	278,646,233	671
http://www.w3.org/2000/01/rdf-schema	58,329,356	533
http://www.w3.org/2002/07/owl	12,934,565	277
http://purl.org/dc/terms/	127,660,537	225
http://www.w3.org/2004/02/skos/core	45,773,412	174
http://purl.org/dc/terms	60,043,900	153
http://xmlns.com/foaf/0.1/	22,121,715	150
http://purl.org/dc/elements/1.1/	9,342,467	141
http://xmlns.com/foaf/0.1	29,489,527	106
http://purl.org/dc/elements/1.1	20,261,128	90
http://rdfs.org/ns/void	12,664	77
http://www.w3.org/2003/01/geo/wgs84_pos	997,603	61
http://www.aktors.org/ontology/portal	33,286,647	54
http://www.w3.org/2006/time	477,210	51
http://www.obofoundry.org/ro/ro.owl	70,167	49
http://creativecommons.org/ns	617,950	47

Ontology usage distribution



Courtesy Li Ding <dingl@cs.rpi.edu>

a. Ontology usage in 2007 as observed by Swoogle

http://dig.csail.mit.edu/2007/Papers/AIMagazine/fractal-paper.pdf

Thing > Person

A person (alive, dead, undead, or fictional).

Property	Expected Type	Description
Properties from Thing		
additionalType	URL	An additional type for the item, typically used for adding more specific types from external vocabularies in microdata syntax. This is a relationship between something and a class that the thing is in. In RDFa syntax, it is better to use the native RDFa syntax - the 'typeof' attribute - for multiple types. Schema.org tools may have only weaker understanding of extra types, in particular those defined externally.
description	Text	A short description of the item.
image	URL	URL of an image of the item.
name	Text	The name of the item.
url	URL	URL of the item.
Properties from Person		
additionalName	Text	An additional name for a Person, can be used for a middle name.
address	PostalAddress	Physical address of the item.
affiliation	Organization	An organization that this person is affiliated with. For example, a school/university, a club, or a team.
alumniOf	EducationalOrganization	An educational organizations that the person is an alumni of.
award	Text	An award won by this person or for this creative work.
awards	Text	Awards won by this person or for this creative work. (legacy spelling; see singular form, award)
birthDate	Date	Date of birth.
brand	Brand or Organization	The brand(s) associated with a product or service, or the brand(s) maintained by an organization or business person.

Re-use and Extensions


I
DC
®

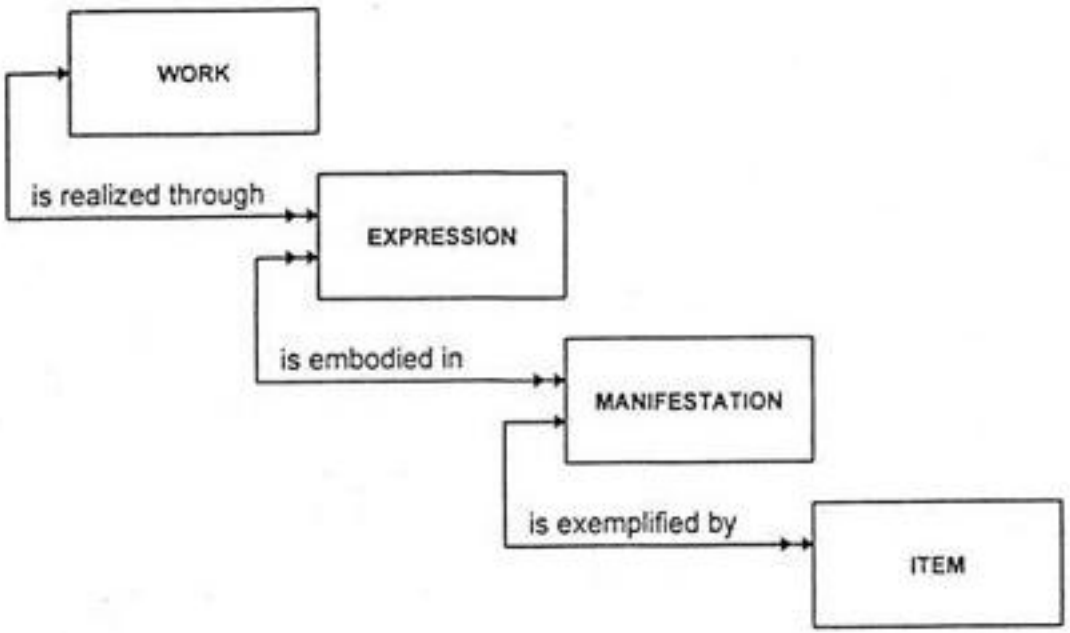
41%

of LOD datasets use Dublin Core

Source: http://lod-cloud.net/state

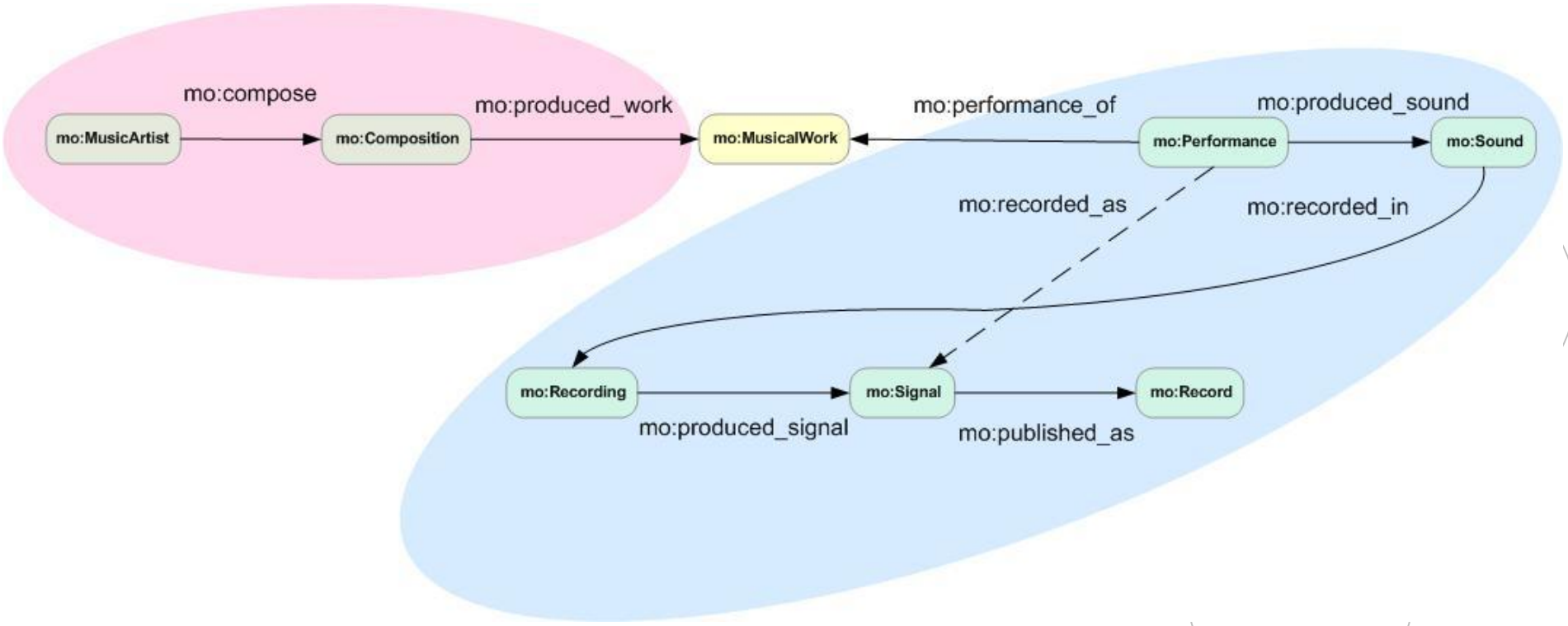

hasPart
type
description
hasFormat
isPartOf
extent
date
replaces
audience
created
requires
issued
conformsTo
coverage
mediator
medium
dateSubmitted
educationLevel
accessRights
rightsHolder
isFormatOf
titleisReferencedBy
accrualPolicy
dateCopyrighted
contributor
bibliographicCitation
dateAccepted
source
spatial
accrualPeriodicity
language
instructionalMethod
isRequiredBy
temporal
valid
tableOfContents
accrualMethod
relation
provenance
publisher
alternative
hasVersion
isVersionOf
available
isReplacedBy
identifier
abstract
references
format
modified
creator
subject
rights
license

FRBR



Source: Jodi Schneider

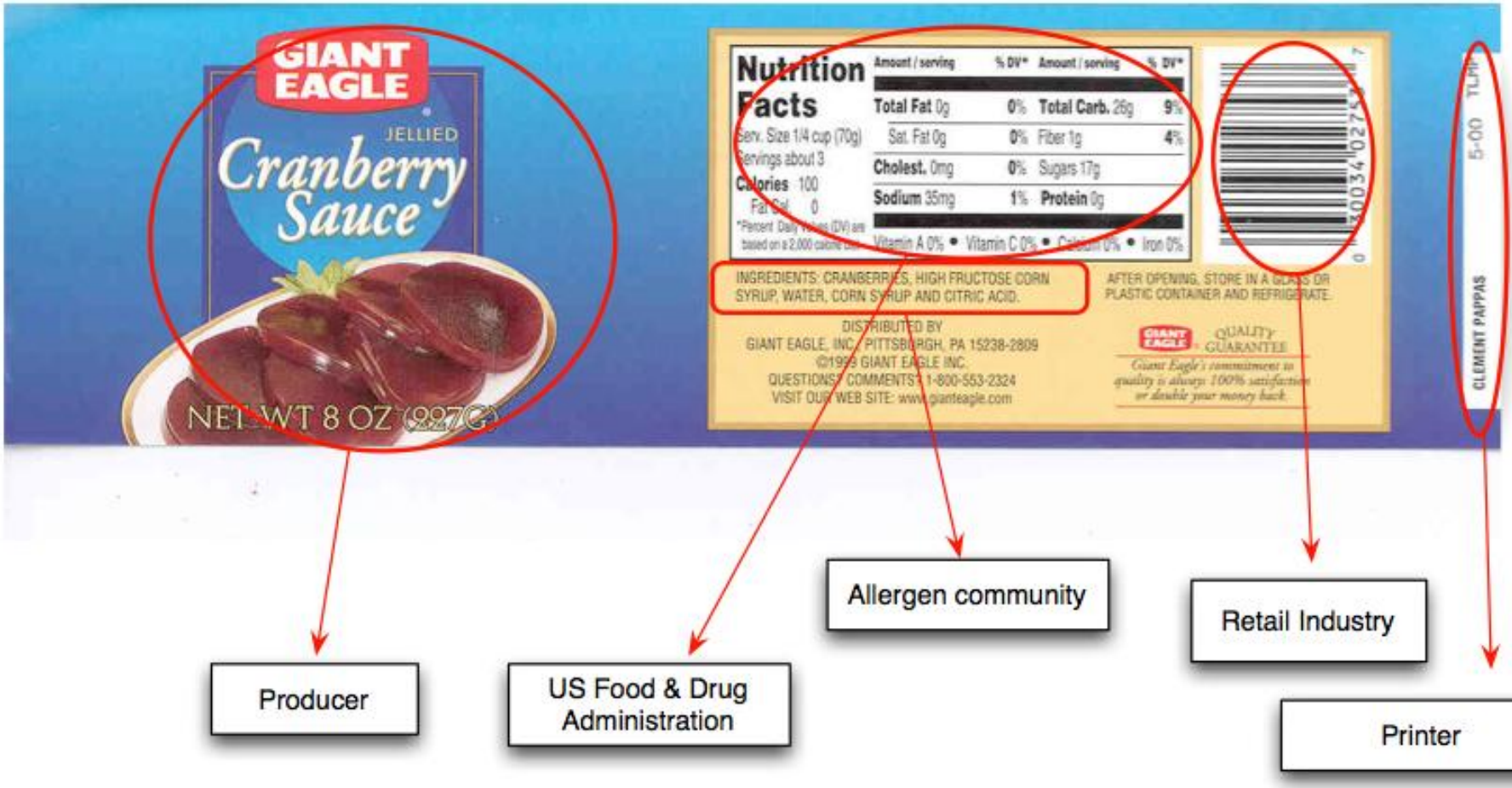
Music Ontology



http://musicontology.com

Theatre Ontology

An ontology for organising theatrical data.



Product label © Giant Eagle, Inc.
http://www.gianteagle.com

http://dig.csail.mit.edu/2007/Papers/AIMagazine/fractal-paper.pdf

Key points

- Ontologies for the web of data tend to be small and simple
- Extension and specialisation for specific communities
- Fractal nature
- Mappings between related ontologies using subClassOf, subPropertyOf, etc.
- Larger cross-community ontologies (schema.org, DBpedia ontology) based on usage evidence