

WHAT'S GOING ON WITH GOVERNMENT LINKED DATA?

CENDI FEDERAL STI MANAGERS GROUP
NASA GODDARD SPACE FLIGHT CENTER
3-NOVEMBER-2011

By: Bernadette Hyland, co-chair
W3C Government Linked Data Working Group

Email: bhyland@3roundstones.com

Twitter: BernHyland

This presentation: <http://slideshare.net/bhylandwood>



“For it is in giving that we receive.”

- Saint Francis of Assisi

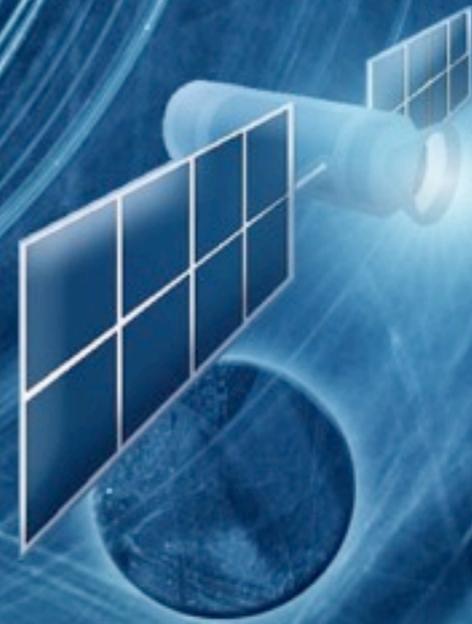
“We make a living by what we do, but we make a life by what we give.”

- Winston Churchill

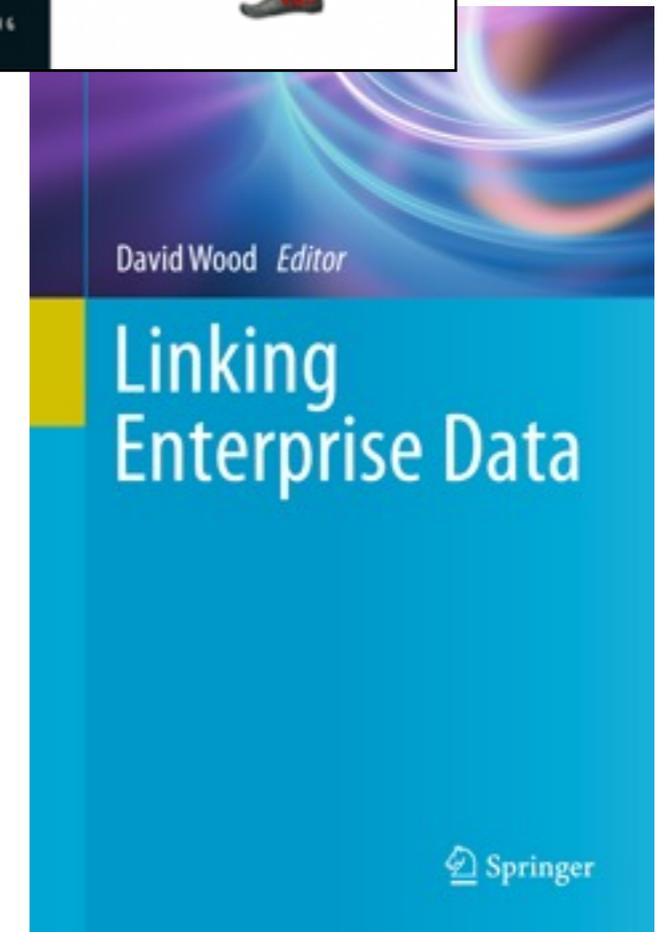
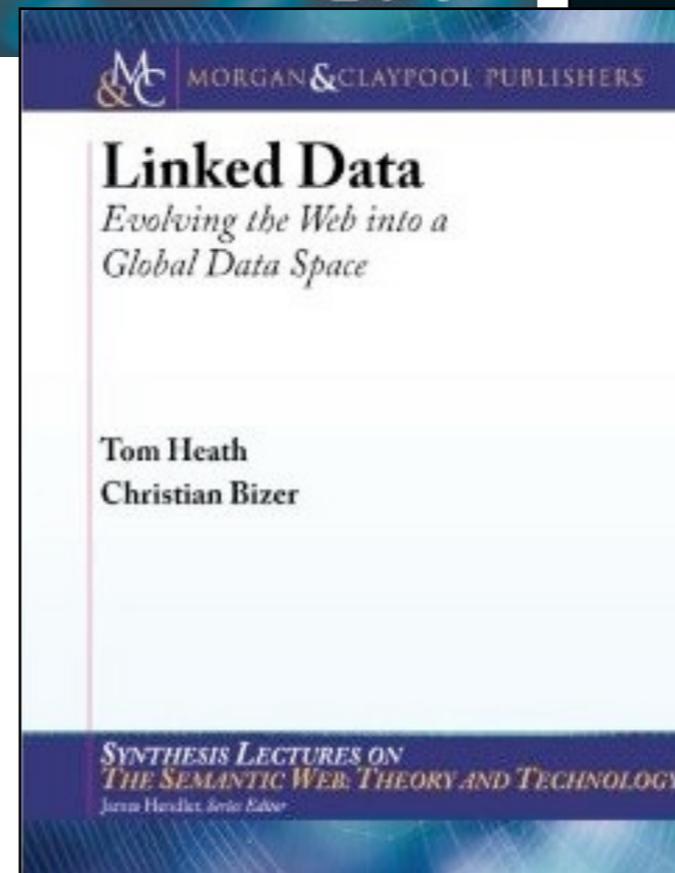
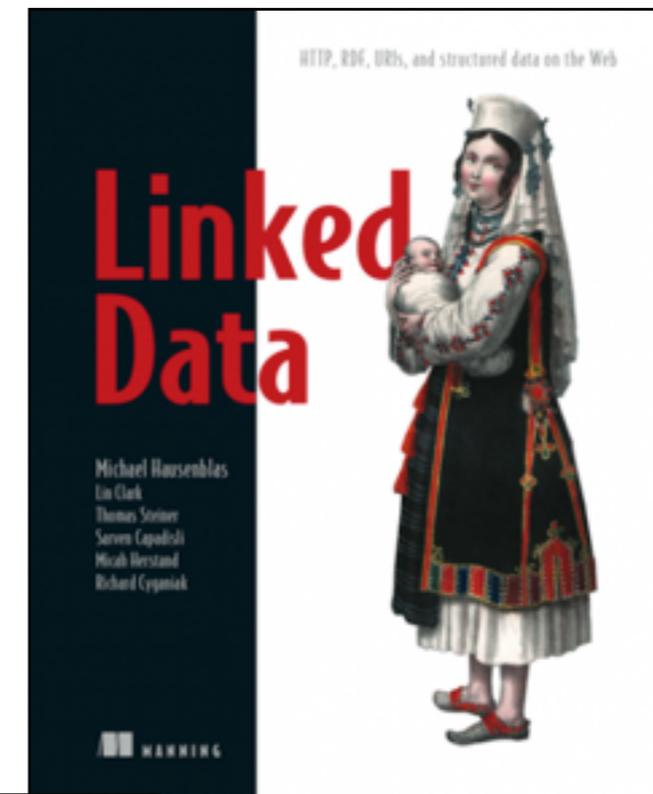
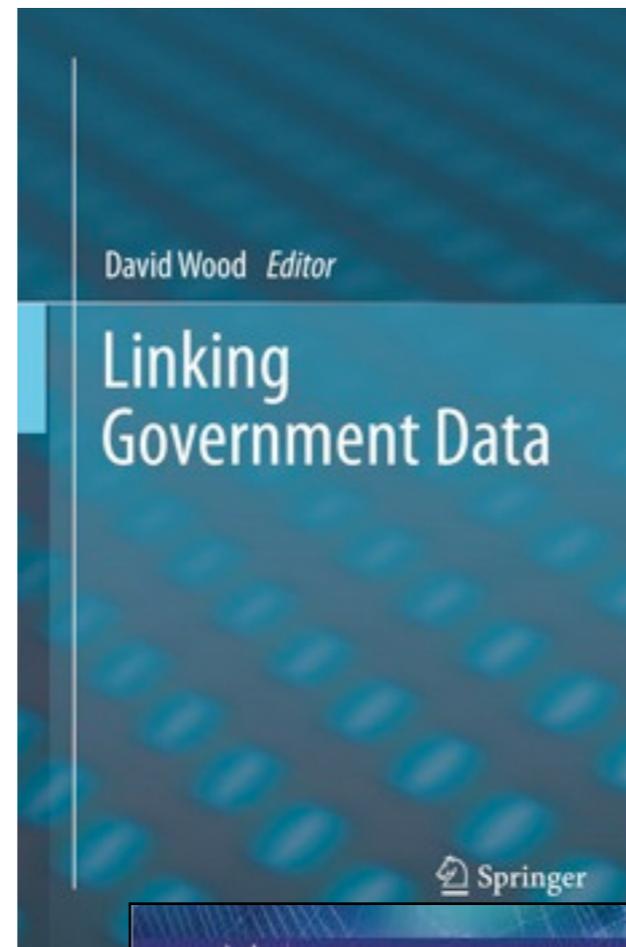
“The miracle is this - the more we share, the more we have.”

- Leonard Nimoy

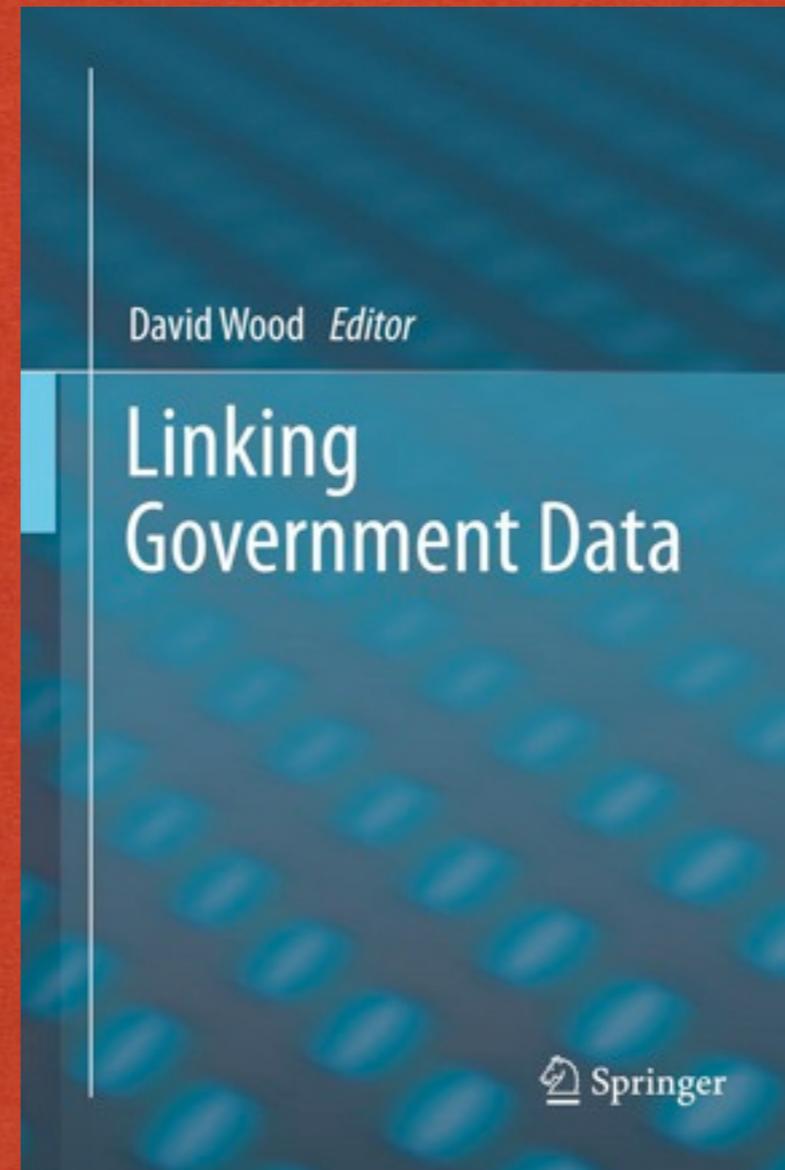
Sharing Worldwide



- Linked Data is about publishing and consuming data using international data standards
- Based on 20 year old idea
- A system of linked information systems

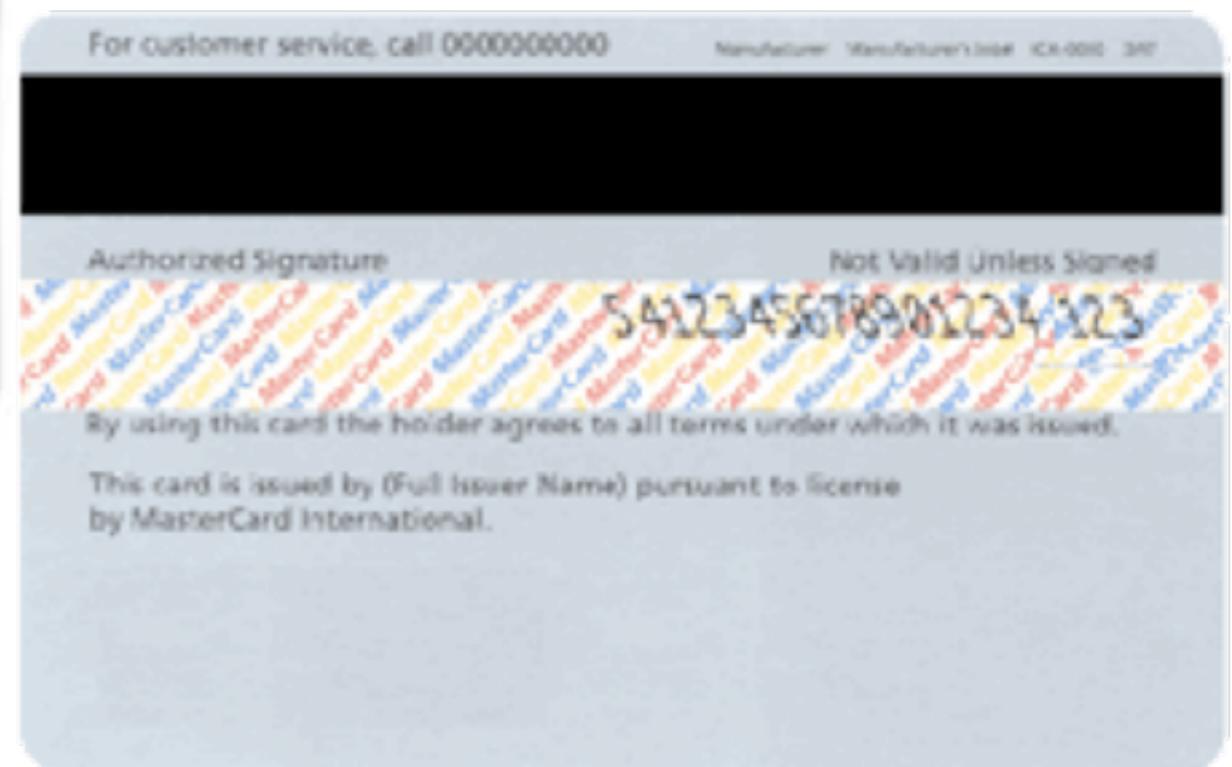


- 42 contributors
- ...from 8 countries
- 10 chapters
- Publication date: November 2011





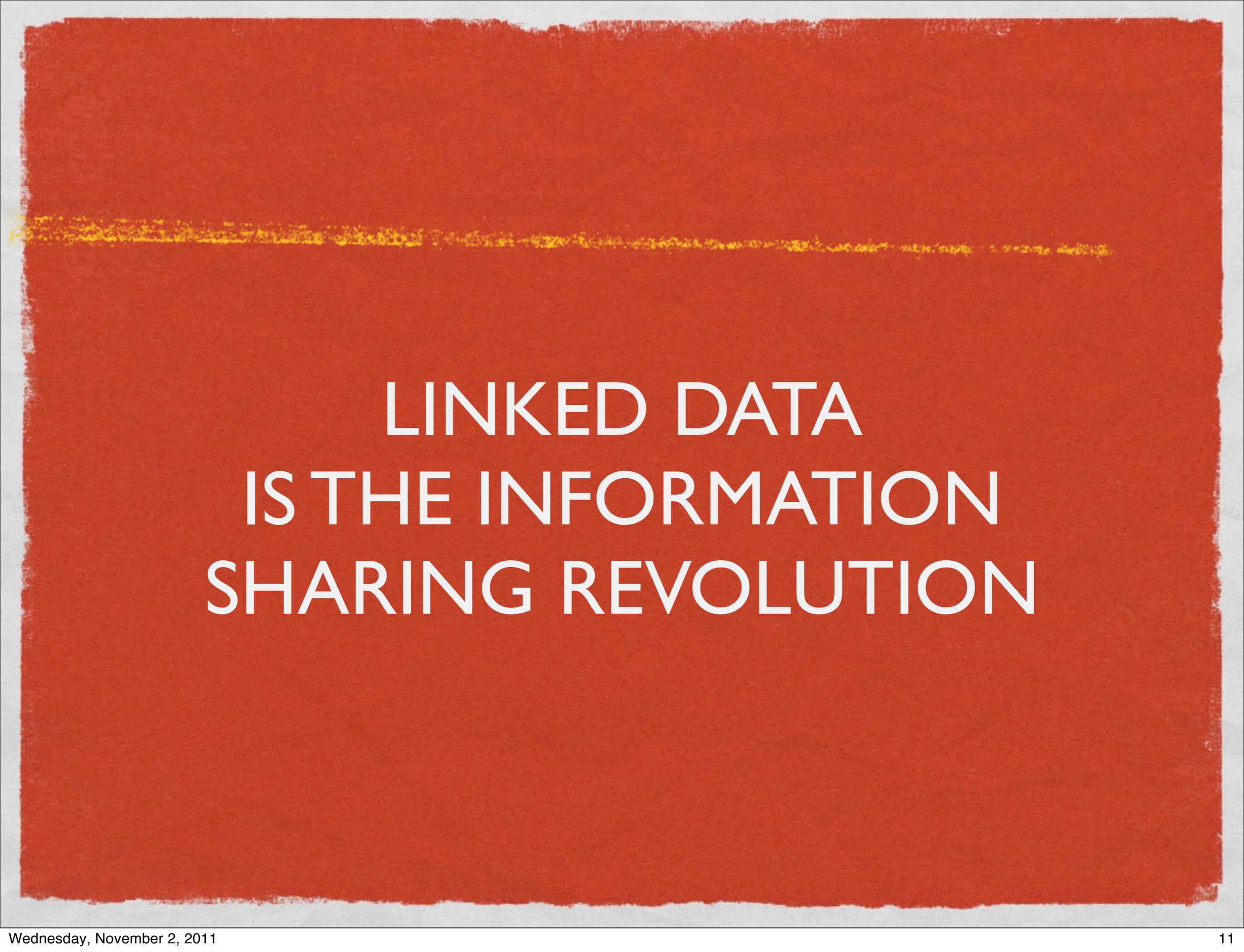
WE'VE SEEN THIS BEFORE





MANY KINDS OF REVOLUTIONS HAPPENING TODAY ...



The background is a solid red color with a textured, brush-painted appearance. A horizontal yellow brushstroke is visible near the top of the page.

LINKED DATA IS THE INFORMATION SHARING REVOLUTION

```

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dc: <http://purl.org/dc/terms/> .
@prefix dctypes: <http://purl.org/dc/dcmitype/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix po: <http://purl.org/ontology/po/> .
@prefix wo: <http://purl.org/ontology/wo/> .

```

```

<http://www.bbc.co.uk/nature/nature/adaptations/Animal_migration#adaptat
a wo:SurvivalStrategy ;
rdfs:label "Migration" .

<http://www.bbc.co.uk/nature/nature/adaptations/Aquatic_locomotion#
a wo:LocomotionAdaptation ;
rdfs:label "Adapted to swimming" .

<http://www.bbc.co.uk/nature/nature/adaptations/Camouflage#ada
a wo:MorphologyAdaptation ;
rdfs:label "Camouflage" .

<http://www.bbc.co.uk/nature/nature/adaptations/Flight#a
a wo:LocomotionAdaptation ;
rdfs:label "Adapted to flying" .

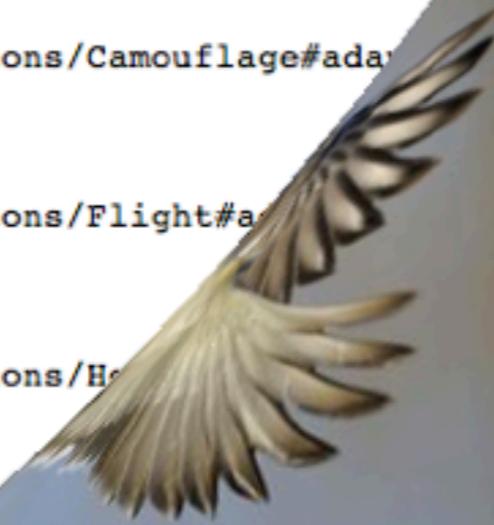
<http://www.bbc.co.uk/nature/nature/adaptations/H
a wo:CommunicationAdaptation ;
rdfs:label "Acoustic communication" .

<http://www.bbc.co.uk/nature/nature/adapt
a wo:LifeCycle ;
rdfs:label "Maternal care" .

<http://www.bbc.co.uk/nature/nat
a wo:LifeCycle ;
rdfs:label "Moulting"

<http://www.bbc.co.uk
a wo:Reproduct
rdfs:label

```



Search the BBC

Mallards are the archetypal and most recognisable ducks. They are found wherever there is water, from rivers and lakes to remote ponds and coastal marshes. It is their adaptability to all kinds of environment that has brought mallards to towns and parks. The males have a glossy green head and white neck ring, but as the females lack in colour they make up for it with noise. The familiar 'quack' is the call of the mother duck summoning her ducklings. Mallards are the ancestors of many breeds of domestic ducks and readily interbreed with them.

Anas platyrhynchos

World duck

DRACILLAM

BBC

Top 10

up a

for a

long leap for fledging

Introduction

Local amenity
 Bill Oddie Goes Wild
 Duck ponds provide a useful resource for wild birds.

Bill Oddie
 Urchfont's domestic ducks are a rowdy lot that still behave like

A mother
 plays a risky game to lure a mink away from

ducklings encounter
 weir for the first time.

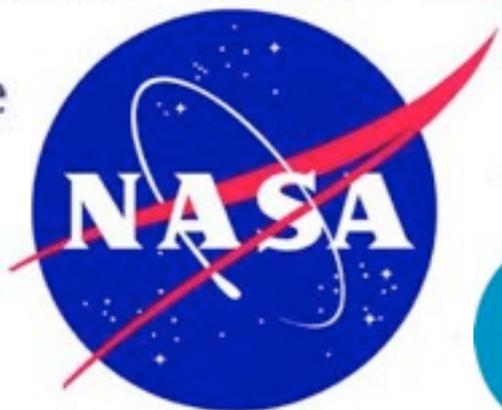
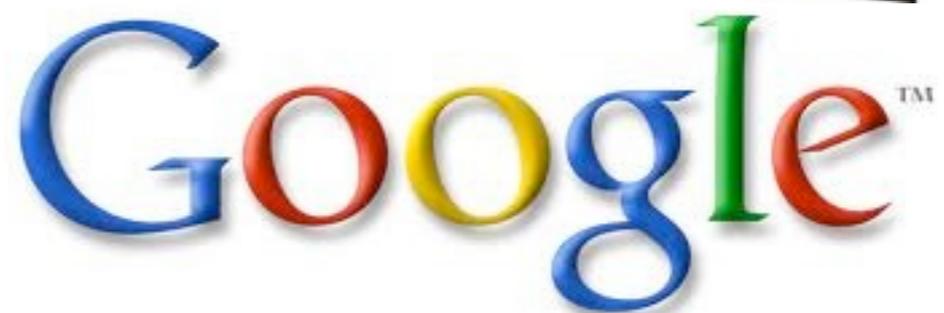
long leap for fledging



ELSEVIER



GOVERNO FEDERAL
BRASIL
PAÍS RICO É PAÍS SEM POBREZA

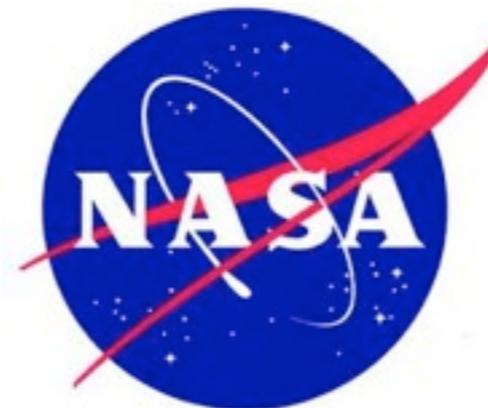


The Open University



Governments

Goals: Governmental transparency and/or improved internal efficiencies (data warehouses)



Hardware/Software Vendors

Goal: Improve interoperability between products and product lines



Retailers

Goal: Improve click-throughs on search results



Book Publishers

Goals: Improve internal manuscript pipelines, expose additional ways of finding and using content



New Media



WHY HAS THERE BEEN GROWTH?

- International data exchange standards are the reason that
 - The Web has become the most extensible, robust information network **ever** created
 - A data sharing revolution is even possible
 - We can share open documents and open data
- Consider the alternative ... **Closed**, proprietary, silos.

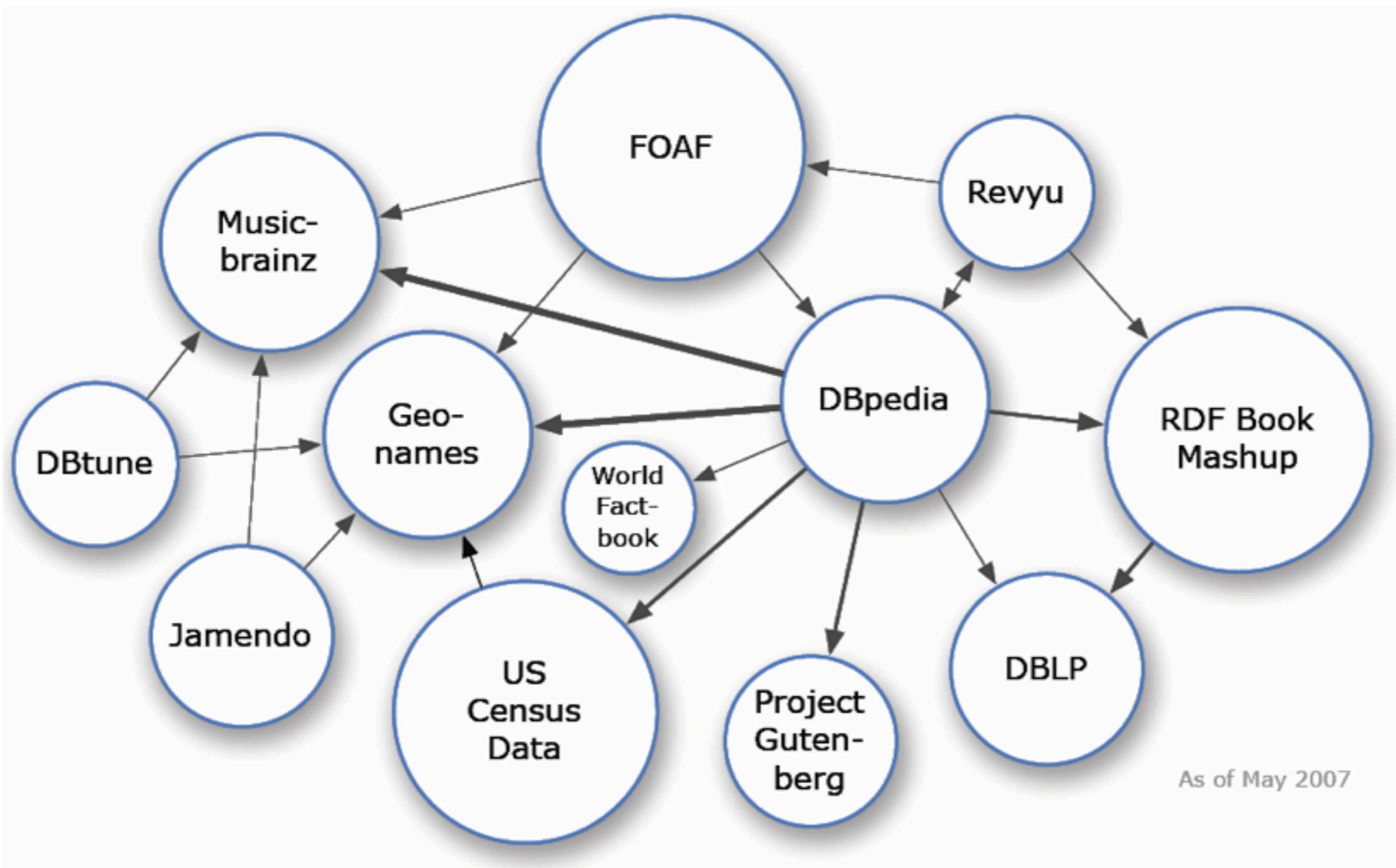
“Cooperation without
coordination”

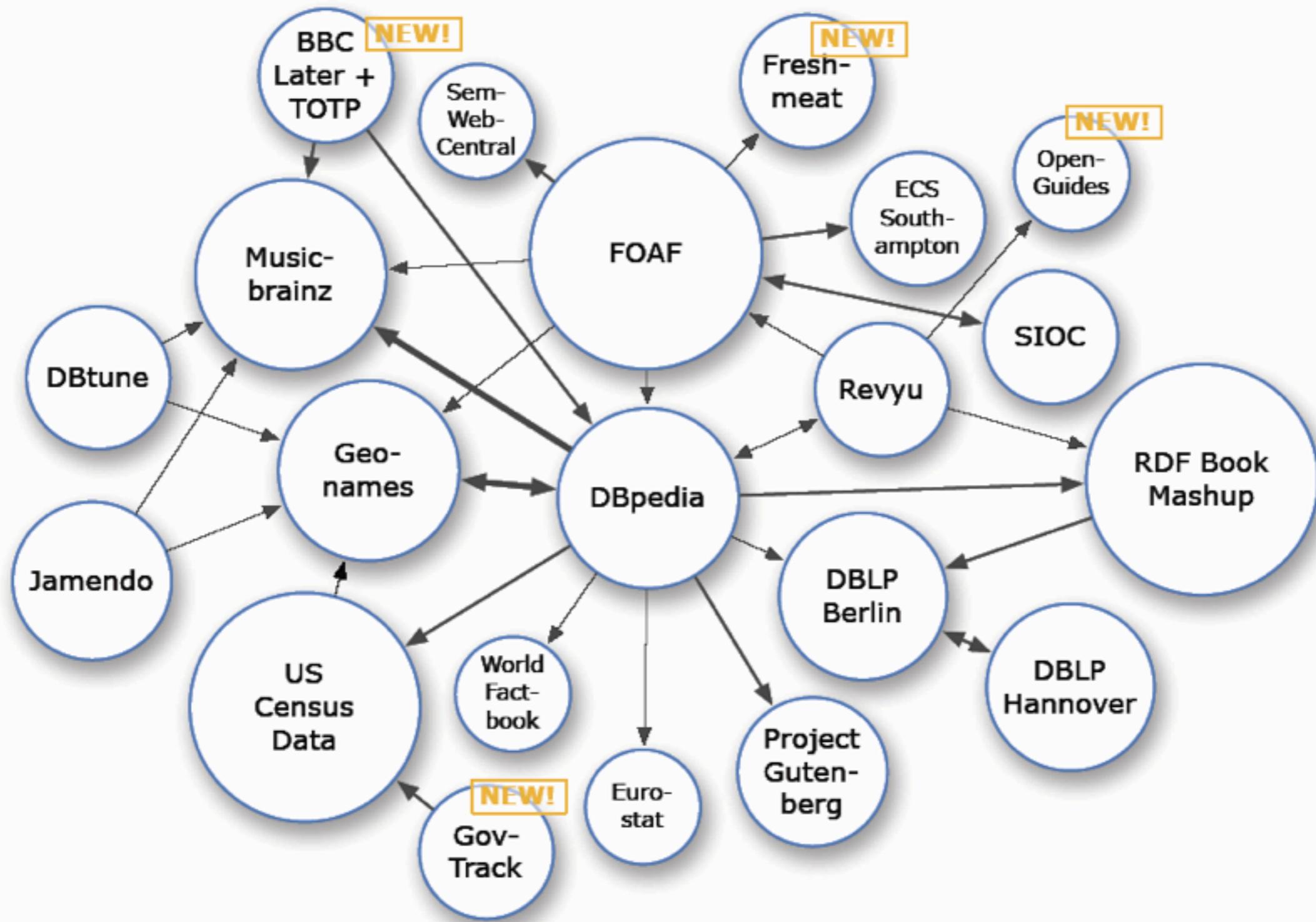
- David Wood

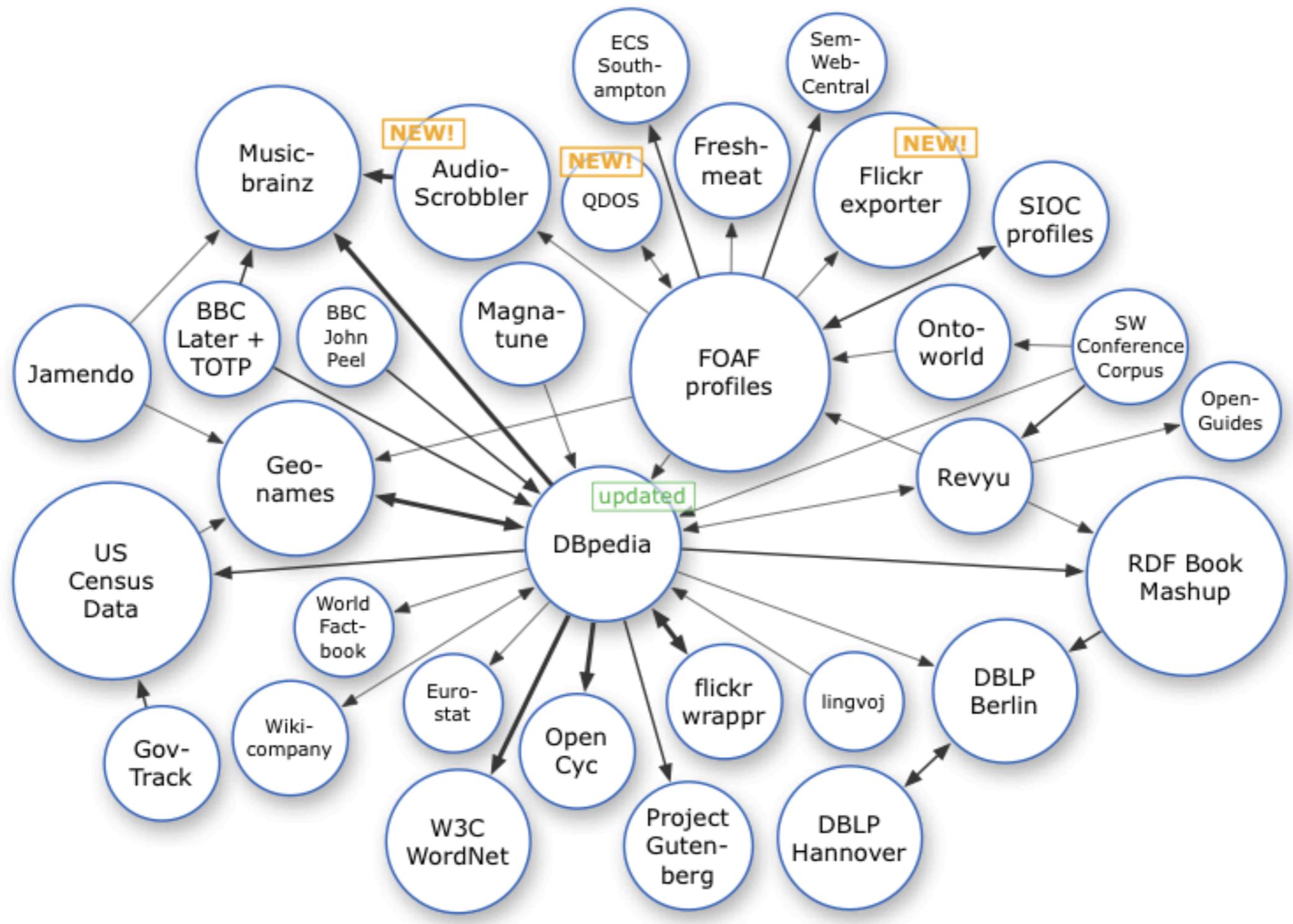
OPEN GOVERNMENT DATA

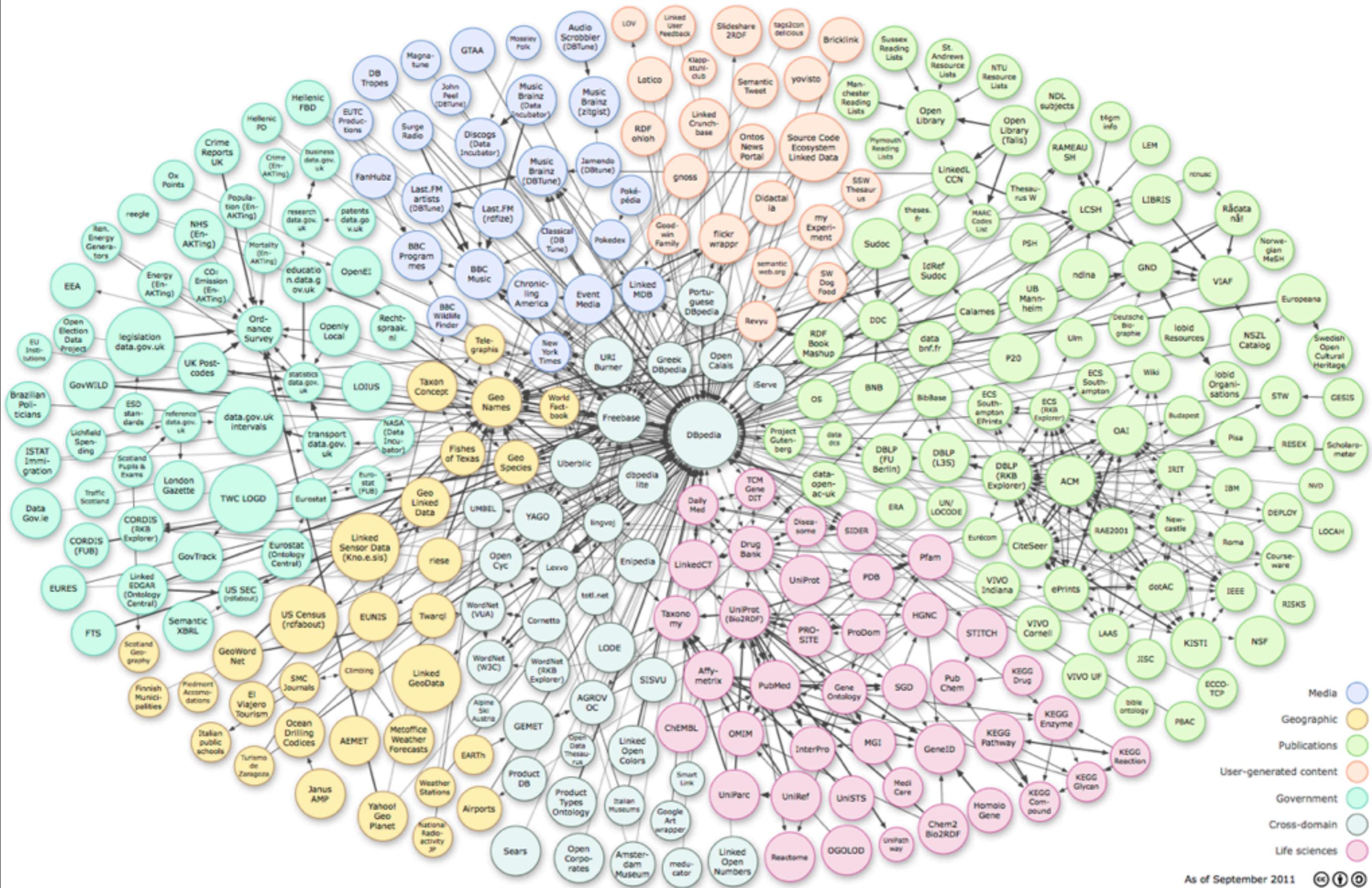
3 BRIEF YEARS ...

- Starting in 2008, a few heads of state directed open government data to be published on the Web
- In September 2011, Presidents Obama (USA) and Rousseff (Brazil) endorsed the Open Government Partnership
- 7 other nations launched their government's National Plans during the meeting of the UN General Assembly



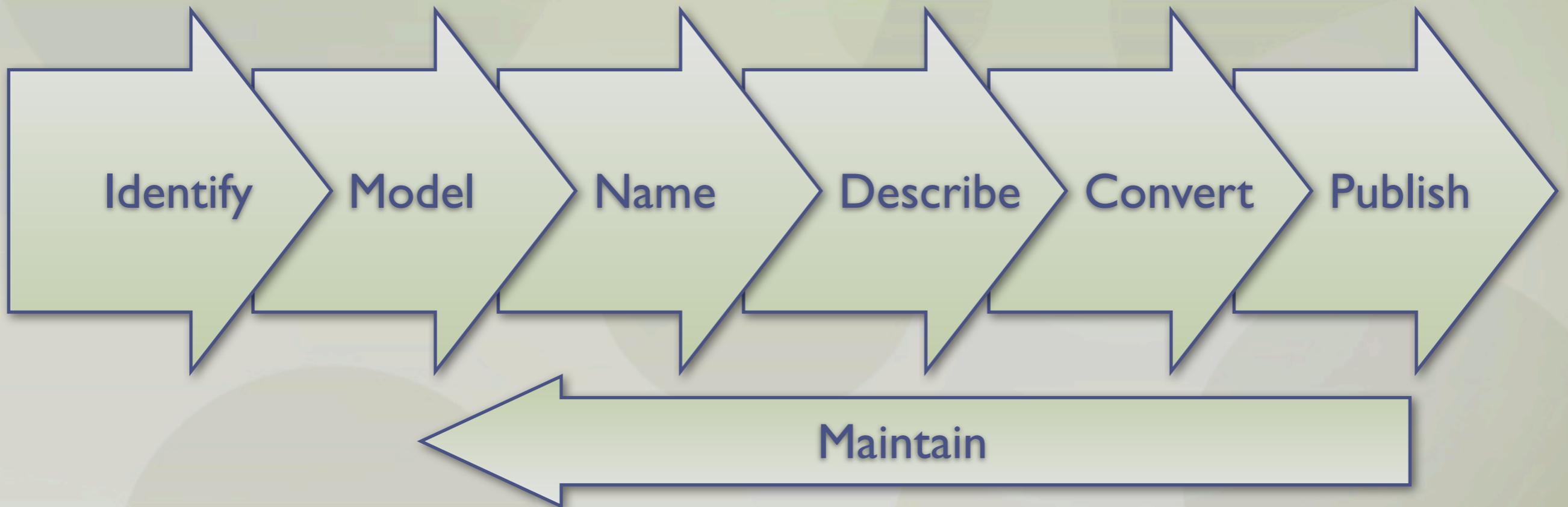




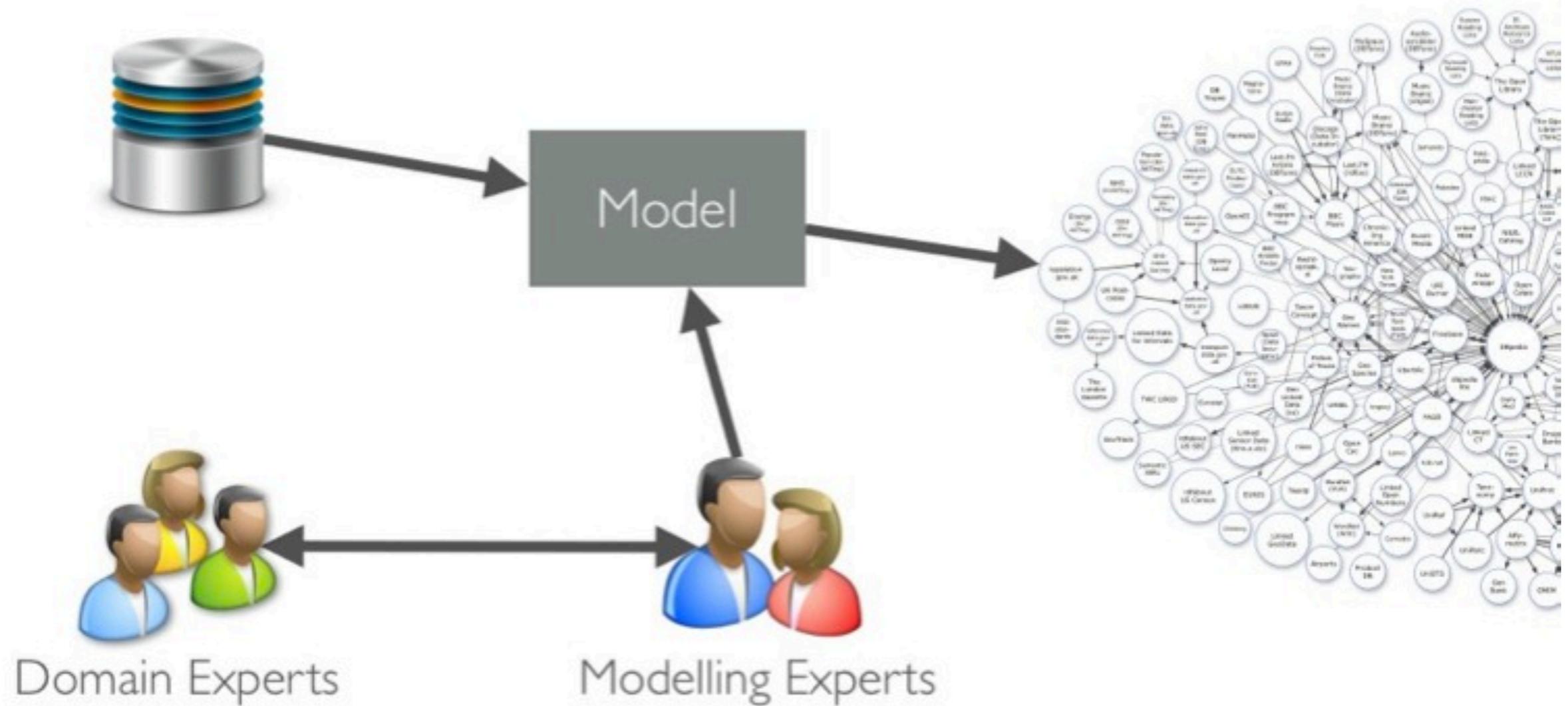




THERE IS A PROCESS



THE PROCESS IS FAMILIAR



A Linked Data modeling approach using Domain and Modeling Expertise

We started with a CSV dump ... ugly, cumbersome data

```
Bernadette-Hyland-Woods-MacBook-2:data bhylandwood$ head NATIONAL_ORGANIZATION_FILE.CSV
"REGISTRY_ID","PGM_SYS_ACRNM","PGM_SYS_ID","INTEREST_TYPE","AFFILIATION_TYPE","START_DATE","END_DATE",
,"DUNS_NUMBER","DIVISION_NAME","PHONE_NUMBER","ALTERNATE_PHONE","FAX_NUMBER","EMAIL_ADDRESS","EIN",
"ENT_NAME","STD_PARENT_NAME","PARENT_DUNS_NUMBER","MAILING_ADDRESS","SUPPLEMENTAL_ADDRESS","CITY_NA",
"AME","POSTAL_CODE","COUNTRY_NAME"
"110000491735","RCRAINFO","AKD000641852","UNSPECIFIED UNIVERSE","OPERATOR","01-APR-04","","FLINT",
"C","PRIVATE","","","9074906217","","","","","","","327 W. SHIP CREEK AVE","","ANCHORAGE",
"110000491735","TRIS","99501MPCLS10760","TRI REPORTER","OWNER/OPERATOR","","","PRIVATE","14330",
,"","","","","","","","",""
"110000491735","RCRAINFO","AKD980987499","SQG","OPERATOR","16-DEC-93","","WILLIAMS ALASKA PETROL",
"","","","","","","","","","","","","",""
"110000491735","OTAQREG","03721","GASOLINE AND DIESEL PRODUCERS","OWNER","","","FLINT HILLS RESOU",
,"","","","","","","4111 E 37TH ST N","","WICHITA","KS","","67220",""
"110000491735","RCRAINFO","AKD980987499","SQG","OWNER","05-JAN-59","","ALASKA RAILROAD CORP","OT",
"","","","327 W. SHIP CREEK AVE","","ANCHORAGE","AK","","99501","USA"
"110000491735","RCRAINFO","AKD000641852","UNSPECIFIED UNIVERSE","OWNER","03-JAN-59","","ALASKA RA",
,"","","","","","","","",""
"110000491735","RCRAINFO","AKD980987499","SQG","OPERATOR","01-APR-04","","FLINT HILLS RESOURCES",
,"9074906217","","",""
"110000491735","TRIS","99501MPCLS10760","TRI REPORTER","PARENT COMPANY","","","KOCH INDUSTRIES IN",
"","","",""
"110000491735","RCRAINFO","AKD000641852","UNSPECIFIED UNIVERSE","OWNER","01-JAN-01","","MAPCO ALA",
"E","","","1076 OCEAN DOCK RD","","ANCHORAGE","AK","","995011199","USA"
Bernadette-Hyland-Woods-MacBook-2:data bhylandwood$
```

Appendix C – EPA-Specific Vocabulary (in Turtle syntax)

```
# Vocabulary for the U.S. Environmental Protection Agency's (EPA) Substance Registry System (SRS) dataset.
```

```
# David Wood (david @ http://3roundstones.com)
```

```
# 6 July 2011
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
```

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
```

```
@prefix owl: <http://www.w3.org/2002/07/owl#> .
```

```
@prefix void: <http://rdfs.org/ns/void#> .
```

```
@prefix srs: <http://us-epa.3roundstones.net/id/us/fed/agency/epa/srs/schema#> .
```

```
<http://us-epa.3roundstones.net/id/us/fed/agency/epa/srs/schema#>
```

```
  a owl:Ontology;
```

```
  rdfs:comment "Vocabulary for the U.S. Environmental Protection Agency's (EPA) Substance Registry System (SRS) dataset." ;
```

```
  void:inDataset <http://us-epa.3roundstones.net/void.ttl#MyDataset> .
```

```
# Classes
```

```
srs:Substance a owl:Class ;
```

```
  rdfs:label "Substance" ;
```

```
  rdfs:comment "A chemical substance, as defined by the US EPA" .
```

```
srs:SubstanceSynonym a owl:Class ;
```

```
  rdfs:label "Substance Synonym" ;
```

```
  rdfs:comment "A synonym for a chemical substance. Synonyms are generally not canonical labels, but have the sense of alternative ways of naming substances." .
```

```
srs:SubstanceList a owl:Class ;
```

```
  rdfs:label "Substance List" ;
```

```
  rdfs:comment "An authority for chemical substance synonyms, such as a database or program that provides such information." .
```


[Main page](#)
[Contents](#)
[All pages](#)
[Recent changes](#)
[Administration](#)
[User groups](#)
[Manifest](#)
[Concept schemes](#)
[Ontologies](#)
[Named queries](#)
[SPARQL query](#)
[Toolbox](#)
[What links here](#)
[Related changes](#)
[Introspect resource](#)
[Permissions](#)
[Print this page](#)
[View](#)
[Edit](#)
[Discussion](#)
[Describe](#)
[History](#)

Styles..

B
I
 x^2
 x_2

H1

Substance Registry System (SRS) Data

P

Search for the names or synonyms of chemical substances in the search field in the upper right to get started.

P

Some sample searches include [Pachynus](#) and [Nitrophenol](#).

P

SRS overview:

- There are 100,036 substances;
- with 548,495 synonyms;
- and 138 substance lists. The most popular sources of synonym records are:
 - [Chemical Abstracts Index Name](#) (75,753)
 - [Chemical Identification](#) (72,939)
 - [TSCA Inventory](#) (66,812)
 - [Synonyms Provided by Submitters to the TSCA Inventory](#) (46,390)
 - [Green Chemistry Expert System](#) (32,101)
 - [Environmental Protection Agency Data Systems](#) (27,075)
 - [Ecotoxicology Database](#) (25,489)
 - [Integrated Taxonomy Information System](#) (17,787)
 - [Storage and Retrieval for Water Quality Data](#) (17,366)
 - [TSCA IUR Exempt](#) (18,610)
 - [elec Inventory Update Rule](#) (14,196)
 - [Radiation Information System for Cleanup Sites](#) (12,046)
- 18,955 substances have kingdom codes, of which 15,813 are in Animalia and 2,877 are in Plantae.
- The SRS data is represented using 4,296,495 RDF statements.

P

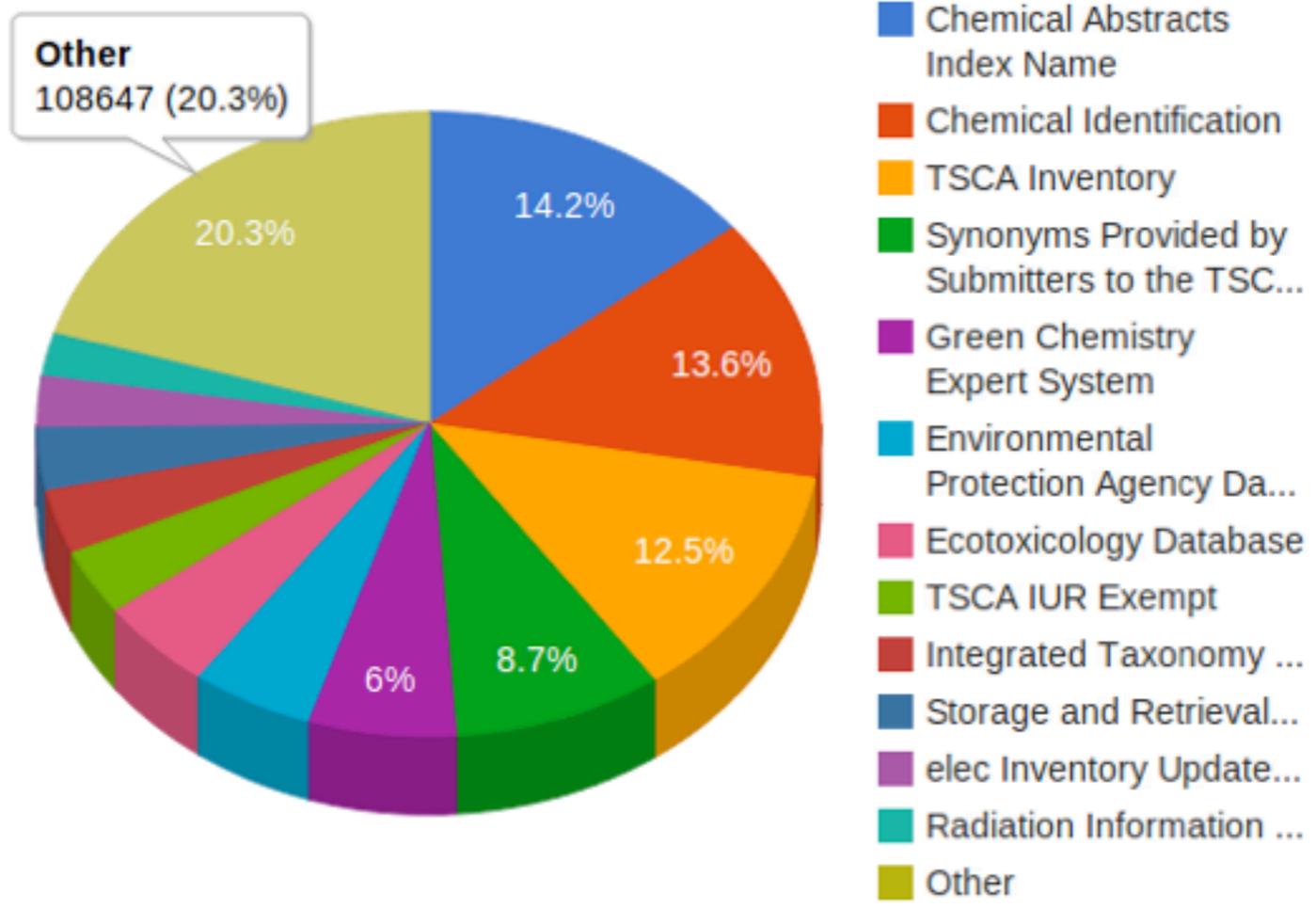
See also some [sample queries](#).

WYMeditor



Number of Synonyms per Hazardous Substance List

- Main page
- Contents
- All pages
- Recent changes
- Administration
 - User groups
 - Manifest
 - Concept schemes
 - Ontologies
 - Named queries
 - SPARQL query
- Toolbox
 - What links here
 - Related changes
 - Introspect resource
 - Permissions
 - Print this page





Administrator | Settings | Contributions

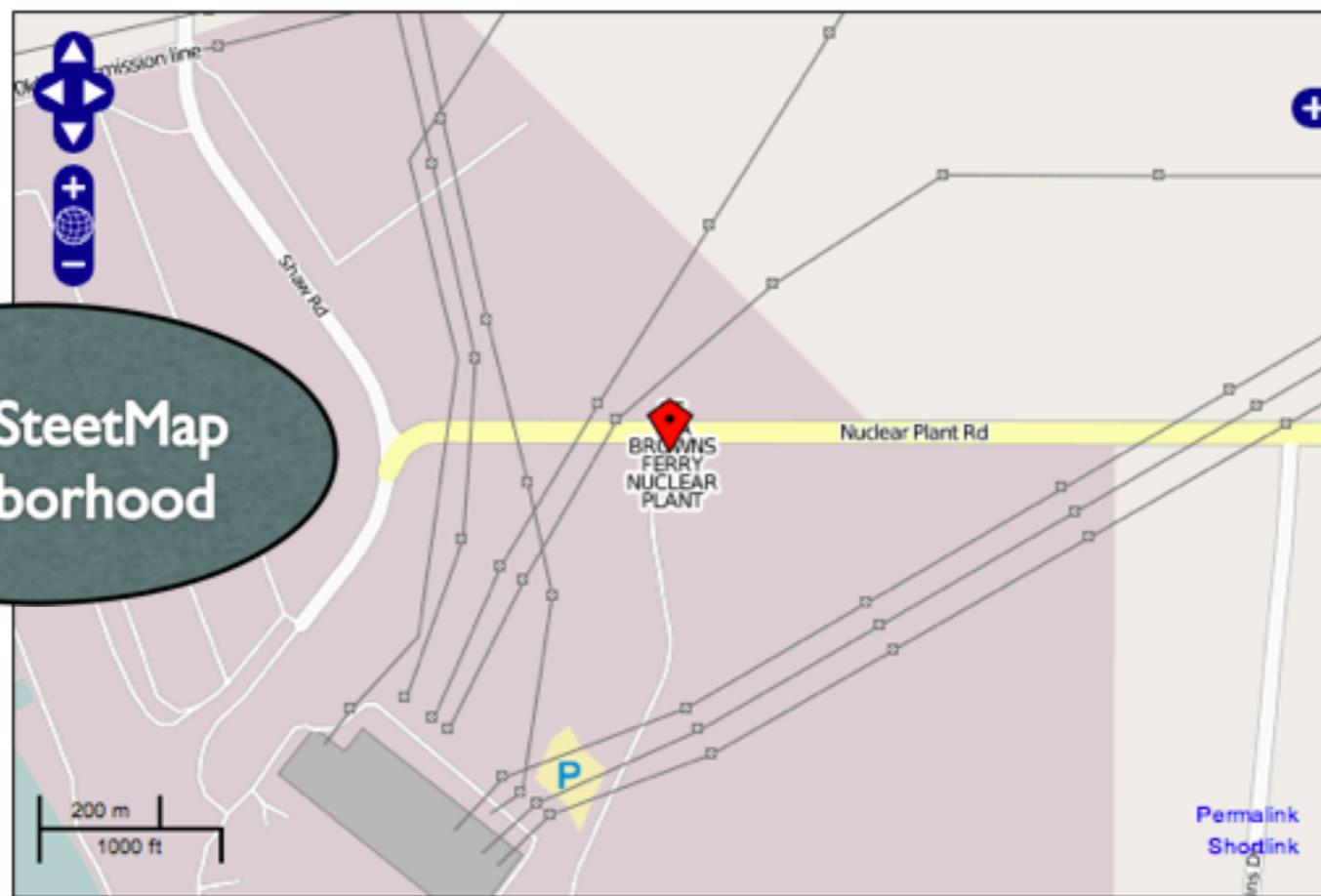
U.s. Tva Browns Ferry Nuclear Plant

Facility ID: 110000589355

OpenStreetMap

log in | sign up

View Edit History Export GPS Traces User Diaries



OpenSteetMap neighborhood

ABT Associates ownership data

Owner

Tennessee Valley Authority (TVA)
◦ Ultimate parent: Tennessee Valley Authority

Operator

Tennessee Valley Authority (TVA)

Street Address

Shaw Rd At Nuclear Plant Rd.
Athens, AL
35611

Mailing Address

Carroll R Cooper
Decatur, Al
35609

FRS

Involved

- Electric Power Generation, Transmission and Distribution
- Nuclear Electric Power Generation
- Electric Power Distribution

Released

- 03lbs of Lead in 2010
- 03lbs of Lead in 2009
- 03lbs of Lead in 2008
- 02lbs of in 2008
- 1lbs of Lead in 2007
- 1lbs of Lead in 2006
- 03lbs of Lead in 2005
- 03lbs of Lead in 2004
- 03lbs of Lead in 2003
- 03lbs of Lead in 2002
- 03lbs of Lead in 2001
- 04lbs of in 2000
- 04lbs of Tetrachloroethylene in 1999
- 04lbs of in 1999
- 04lbs of Tetrachloroethylene in 1998
- 04lbs of in 1998

TRI & SRS



Administrator | Settings | Contributions | Logout

U.S. TVA Browns Ferry Nuclear Plant

Name

Fix name

Comment

Add a comment

[View](#) | [Edit](#) | [Discussion](#) | [Describe](#) | [History](#)

Site

- Main article
- Home folder
- Contents
- Recent changes

Resources

- Folders
- Articles
- Pages
- Concepts
- Classes

Administration

- Manifest
- Users
- Groups
- Named queries
- SPARQL query

Toolbox

- What links here
- Related changes
- Introspect resource
- Permissions
- Print this page



This resource was last modified at 10:10 pm



Profile | Settings | Contributions | Logout

Tennessee Valley Authority



The Tennessee Valley Authority (TVA) is a federally owned corporation in the United States created by congressional charter in May 1933 to provide navigation, flood control, electricity generation, fertilizer manufacturing, and economic development in the Tennessee Valley, a region particularly affected by the Great Depression. The enterprise was a result of the efforts of Senator George W. Norris of Nebraska. TVA was envisioned not only as a provider, but also as a regional economic development agency that would use federal experts and electricity to rapidly modernize the region's economy and society. TVA's service area covers most of Tennessee, parts of Alabama, Mississippi, and Kentucky, and small slices of Georgia, North Carolina, and Virginia. It was the first large regional planning agency of the federal government and remains the largest. Under the leadership of David Lilienthal ("Mr. TVA"), TVA became a model for America's governmental efforts to modernize Third World agrarian societies.

View | Edit | Discussion | Describe | History |

SEC Identifier
0001376986

Web Resources

- [Home page](#)
- [Wikipedia entry](#)

Ultimate parent

- [Tennessee Valley Authority](#)

Operator of

- [U.S. TVA Browns Ferry Nuclear Plant](#)
- [U.S. TVA Sequoyah Nuclear Plant](#)
- [U.S. TVA Watts Bar Nuclear Plant](#)

Site

Main article
Home folder
Contents
Recent changes

3 Round Stones

[Linked Data Demo](#)

Resources

Folders
Articles
Text files
Pages
Scripts
Styles
Concepts
Classes

Administration

Manifest
User groups
Named queries
SPARQL query

Toolbox

What links here
Related changes
Introspect resource
Permissions
Print this page



This resource was last modified at 8:28 pm, 28 Sep

U.S. TVA Browns Ferry Nuclear Plant

Facility ID: 110000589355

The Browns Ferry Nuclear Plant is located on the Tennessee River near Decatur [...]

 OpenStreetMap

[log in](#) | [sign up](#)

[View](#) [Edit](#) [History](#) [Export](#) [GPS Traces](#) [User Diaries](#)



Owner

Tennessee Valley Authority (TVA)

- **Ultimate parent** Tennessee Valley Authority

Operator

Tennessee Valley Authority (TVA)

Street Address

Carroll R Cooper
Decatur, AL
35609

Mailing Address

Carroll R Cooper
Decatur, AL
35609

Involved

- Electric Power Generation, Transmission and Distribution
- Nuclear Electric Power Generation
- Electric Power Distribution

Released:

- Lead in 2010
- Lead in 2009
- Lead in 2008
- Polycyclic aromatic compounds in 2008
- Lead in 2007
- Lead in 2006
- Lead in 2005
- Lead in 2004
- Lead in 2003
- Lead in 2002
- Lead in 2001
- Zinc compounds in 2000
- Tetrachloroethylene in 1999
- Zinc compounds in 1999
- Tetrachloroethylene in 1998
- Zinc compounds in 1998

LINKED DATA PLATFORM

- Callimachus is a framework for data-driven applications based on [Linked Data](#) principles
- Callimachus allows Web developers to easily create data driven applications for the Web
- <http://callimachusproject.org>



Callimachus

THE SOCIAL CONTRACT ...

THE NOT SO FINE PRINT

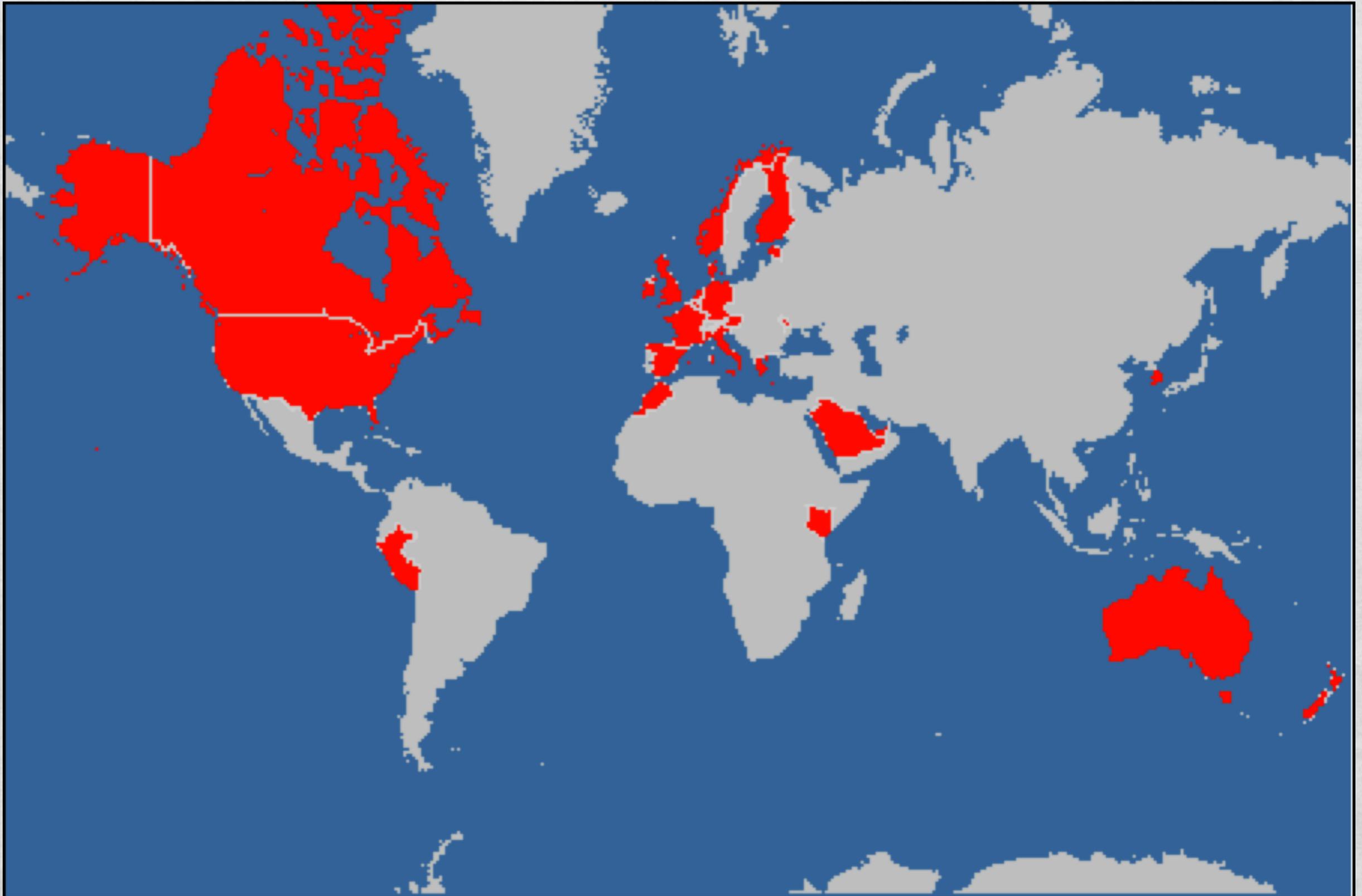
- Linked Open Data is a *social contract* to provide information
- Ensure that your Linked Open Data remains available
- Follow best practices for modeling
- Carefully consider your URI strategy
- Publish a VoID description
- For a government agency ... a **data policy is “a must”**
 - specify data quality and retention, treatment of data thru secondary sources, restrictions for use, frequency of updates, public participation, and applicability of this data policy



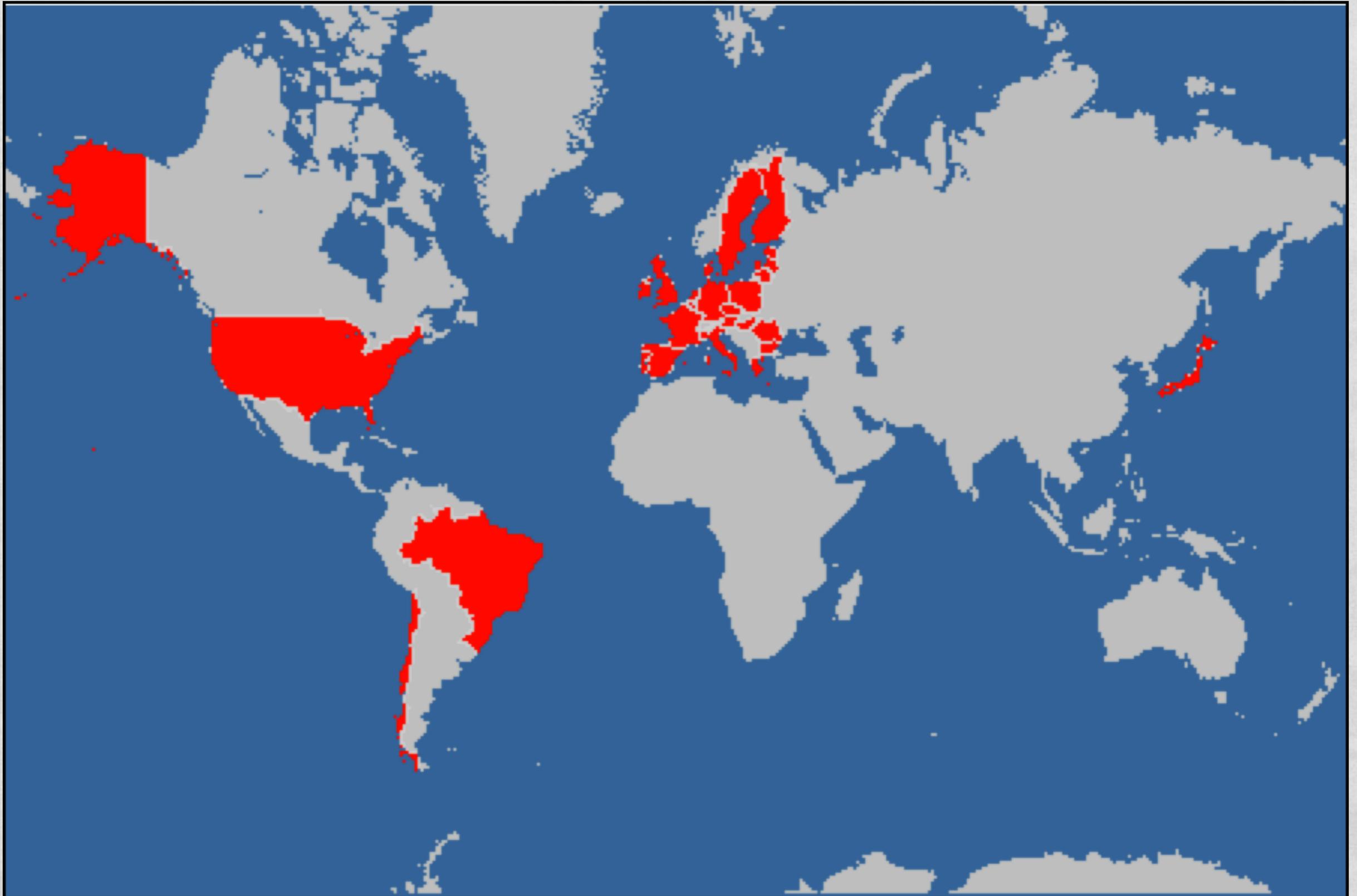
Open Government Data Camp 2011



Open Government Data Sites in 2011



Government LOD on CKAN



Largest Publisher of Government Linked Open Data



HOW DO WE START?

Where does one get training and mentoring?

Is there guidance for my procurement department?

How do we define the value proposition & ROI?

What is the Government Policy? FISMA?
Authority to Operate?

W3C GOVERNMENT LINKED DATA WORKING GROUP



- Started June 2011; runs to May 2013
- Chartered to provide standards & develop standards track documents to help all governments share their data as high quality (“5 star”) Linked Data
- 39 participants from 25 organizations
- 50% in non-US locations
- <http://www.w3.org/2011/gld/charter.html>

DELIVERABLES

- Community Directory
- Best Practices for Publishing Linked Data
 - Procurement, vocabulary selection, URI construction, versioning, stability, legacy data issues
- Cookbook for Linked Open Data
- Standard Vocabularies
 - Metadata, Statistical “Cube” Data, People, Organizational structures



- Scope
- Deliverables
- Dependencies and Liaisons
- Participation**
- Communication
- Decision Policy**
- Patent Policy**
- About this Charter**

Government Linked Data Working Group Charter

The **mission** of the [Government Linked Data \(GLD\) Working Group](#) is to provide standards and other information which help governments around the world publish their data as effective and usable [Linked Data](#) using [Semantic Web](#) technologies. The group, a part of the [eGovernment Activity](#) and closely connected with the [Semantic Web Activity](#), will collect and make available information about government Linked Data activities around the world. It will use that information and the experience of its participants to develop W3C Recommendations for Best Practices and for RDF Vocabularies necessary for publication of government data in RDF, as Linked Data. Concurrently with this Working Group, W3C has chartered the [eGovernment Interest Group](#) for broad community discussion of government use of the Web

[How to Join the GLD Working Group](#)

Start Date	June 2011
End date	31 May 2013
Confidentiality	Proceedings are public
Chairs	George Thomas, US HHS Bernadette Hyland
Staff Contact	Sandro Hawke (0.3 FTE)
Usual Meeting Schedule	Teleconferences: Weekly, plus optional task forces Face-to-face: 2-4 per year, possibly multisite with video

<http://www.w3.org/2011/gld/charter>



ORGANIZATIONS BY CATEGORY

10 Items filtered from 29 originally (Reset All Filters)

Organization ▲
3 Round Stones Inc.
Centers for Medicare and Medicaid Services
CTIC
Gijón City Council
Government of Catalonia
International Business Machines
netage.nl
Ordnance Survey
U.S. Department of Health & Human Services
United States Environmental Protection Agency

List all entries

- Category** 1
- 5
 - 8 Consulting
 - 10 **Government - Federal or Central**
 - 1 Government - Local
 - 4 Middleware
 - 11 Ontology Development

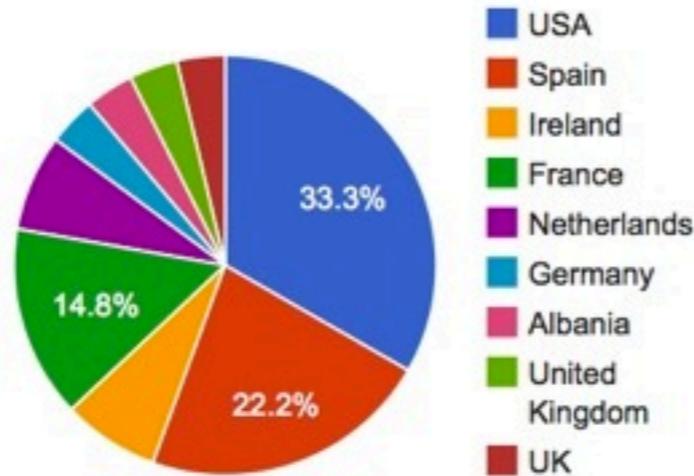
- Country**
- 1 Netherlands
 - 3 Spain
 - 1 UK
 - 5 USA

- Organization Type**
- 3 Commercial
 - 4 Federal / Central Govt
 - 2 Local Government
 - 1 Research / Academic

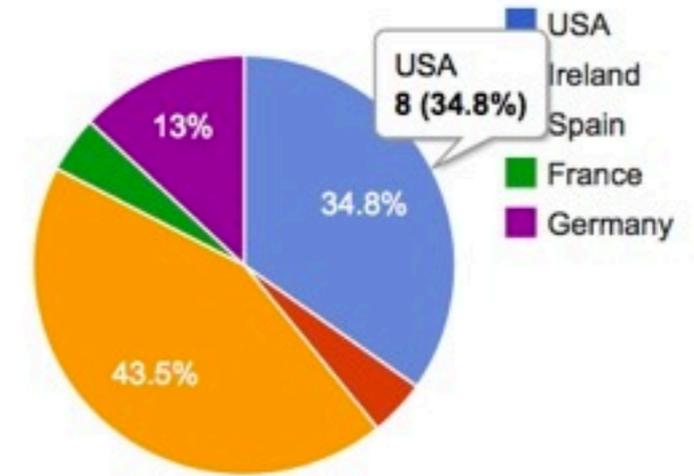
- Community Directory**
- About
- Directory
- View by Category
- View statistics
- Site**
- Main page
- Contents
- All pages
- Recent changes
- Administration**
- User groups
- Concept schemes
- Named queries
- SPARQL query
- Toolbox**
- What links here
- Permissions
- Print this page
- Related changes

STATISTICS

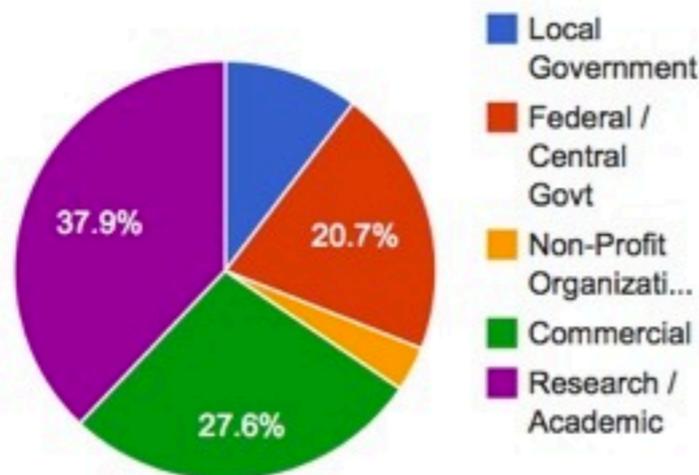
Organizations by Country



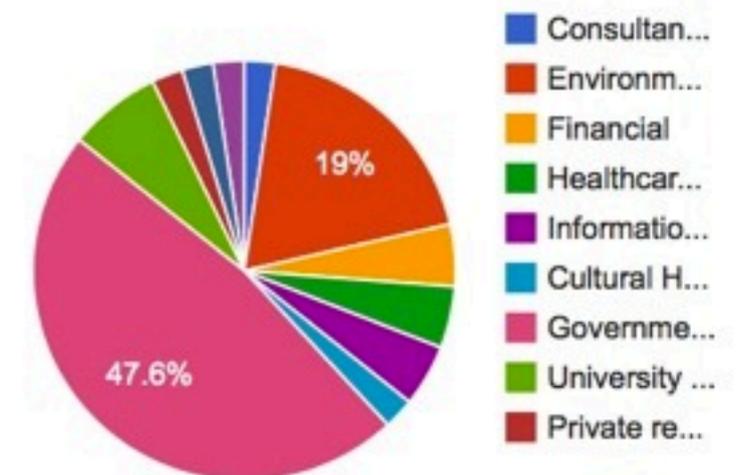
Projects by Country



Projects by Type of Organization

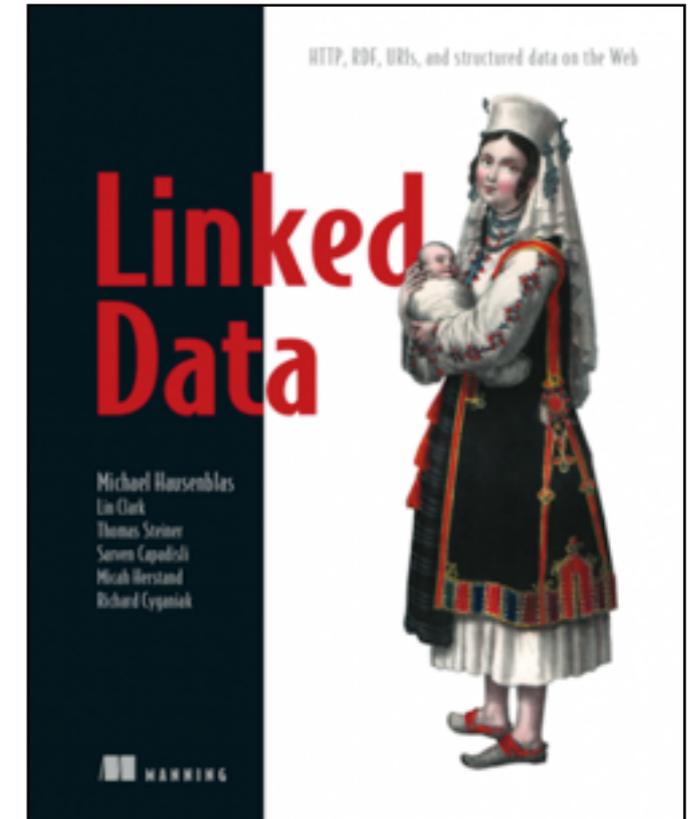
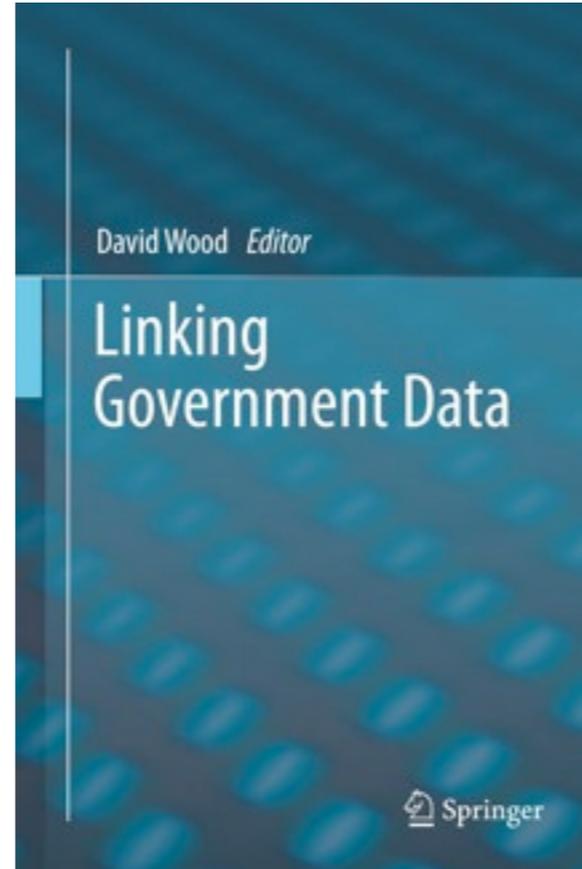
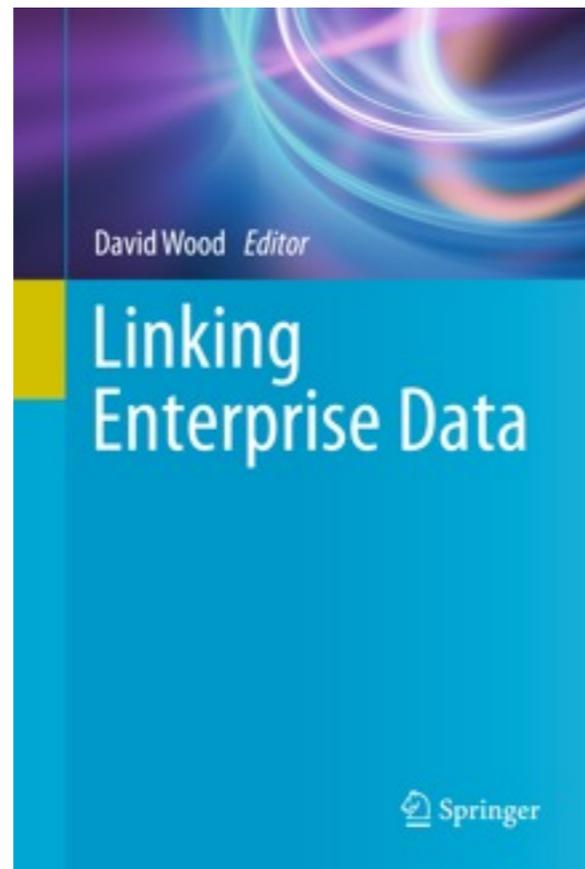
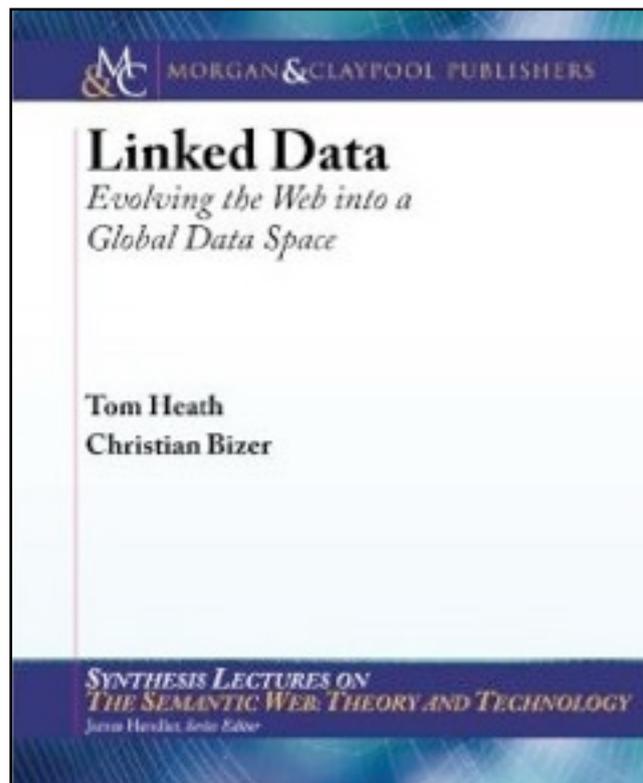


Projects by Industry or Domain Category



- [Community Directory](#)
- [About](#)
- [Directory](#)
- [View by Category](#)
- [View statistics](#)
- Site**
- [Main page](#)
- [Contents](#)
- [All pages](#)
- [Recent changes](#)
- Administration**
- [User groups](#)
- [Concept schemes](#)
- [Named queries](#)
- [SPARQL query](#)
- Toolbox**
- [What links here](#)
- [Permissions](#)
- [Print this page](#)
- [Related changes](#)

Further Reading



<http://linkeddatabook.com/editions/1.0/>

<http://3roundstones.com/linking-enterprise-data/>

<http://3roundstones.com/linking-government-data/>

<http://www.linkeddatadeveloper.com/>

THANK YOU FOR THE OPPORTUNITY TO PARTICIPATE

Bernadette Hyland

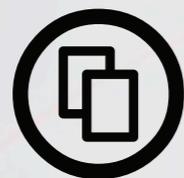
Co-chair W3C Government Linked Data Working Group & CEO 3 Round Stones, Inc

- Twitter: @BernHyland
bhyland@3roundstones.com
- <http://slideshare.net/bhylandwood>



This work is Copyright © 2011 3 Round Stones Inc.
It is licensed under the Creative Commons Attribution 3.0 Unported License
Full details at: <http://creativecommons.org/licenses/by/3.0/>

You are free:



to Share — to copy, distribute and transmit the work



to Remix — to adapt the work

Under the following conditions:



Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

Credits

Riga Marathon, By Nordea (CC-BY-NC)	http://www.flickr.com/photos/nordearigamarathon/4549346580/
Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara (public domain)	http://proton.ucting.udg.mx/galeria/3D/WEB.jpg
I-5 Star Linked Data image	http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/
LOD Cloud Diagrams Richard Cyganiak, Anja Jentzsch, (CC-BY-SA)	http://lod-cloud.net/
Arab uprising photo, Collin David Anderson (CC-BY-NC- SA)	http://www.flickr.com/photos/collina/5459554667/
Book covers © their respective owners and used under Fair Use for educational purposes	
All other photos © 2011 Bernadette Hyland, released under a CC-BY-SA license	



Credits

Head with 3D Glasses Mark Chapman CC-BY licensed	http://www.flickr.com/photos/markchapmanphoto/5139429152/
Chasm Photo Travis S. CC-BY-NC licensed	http://www.flickr.com/photos/baggis/3860802929/
Occupy Wall Street by emilydickinsonridesabmx CC-BY licensed	http://www.flickr.com/photos/emilyrides/6212739587/sizes/l/in/photostream/
Sharing Squirrels Image leezie5 CC-BY-NC-ND licensed	http://www.flickr.com/photos/leeziet/5912219625/