

- Watch the web site for schedule (especially wrt Nov 20)!
- Read the paper on „web science“ as indicated on
<http://west.uni-koblenz.de/teaching/ws1213/webscience>
it will be considered part of the curriculum and will be asked for in the oral exam

Structured Data on the Web

Introduction to Web Science

Steffen Staab

Slides from Jim Hendler & Semantic Web Company



- HTML Meta
- Microdata
- Microformats
- RDF
 - ◆ RDF-a
- Applications
 - ◆ Schema.org – done already
 - ◆ Wikidata - seminar
 - ◆ Open data (especially Open government data)

OGD Austria - <http://gov.opendata.at>

Open Government Data is a worldwide movement to open data (& information) of the government / public administration* - that is NOT personal (individual related) – in human- AND machine-readable open (non proprietary) formats for re-use (by civil society, economy, media, academia AND politics & public administration itself)!

OPEN stands for lowering the barriers to ensure as broad as possible re-use (for everybody)!

There is a **new paradigm** in publishing Open Government Data
= look, take and play!

* data and information produced or commissioned by government or government controlled entities

1. Completeness

All public data are made available. Public data are data that are not subject to valid privacy, security or privilege limitations, as governed by other statutes.

2. Primacy

Data are published as collected at the source, with the finest possible level of granularity, not in aggregate or modified forms.

3. Timeliness

Data are made available as quickly as necessary to preserve the value of the data.

4. Ease of Physical and Electronic Access

Data are available to the widest range of users for the widest range of purposes.

5. Machine readability

Data are reasonably structured to allow automated processing of it.

6. Non-discrimination

Data are available to anyone, with no requirement of registration.

7. Use of Commonly Owned Standards (e.g. open formats)

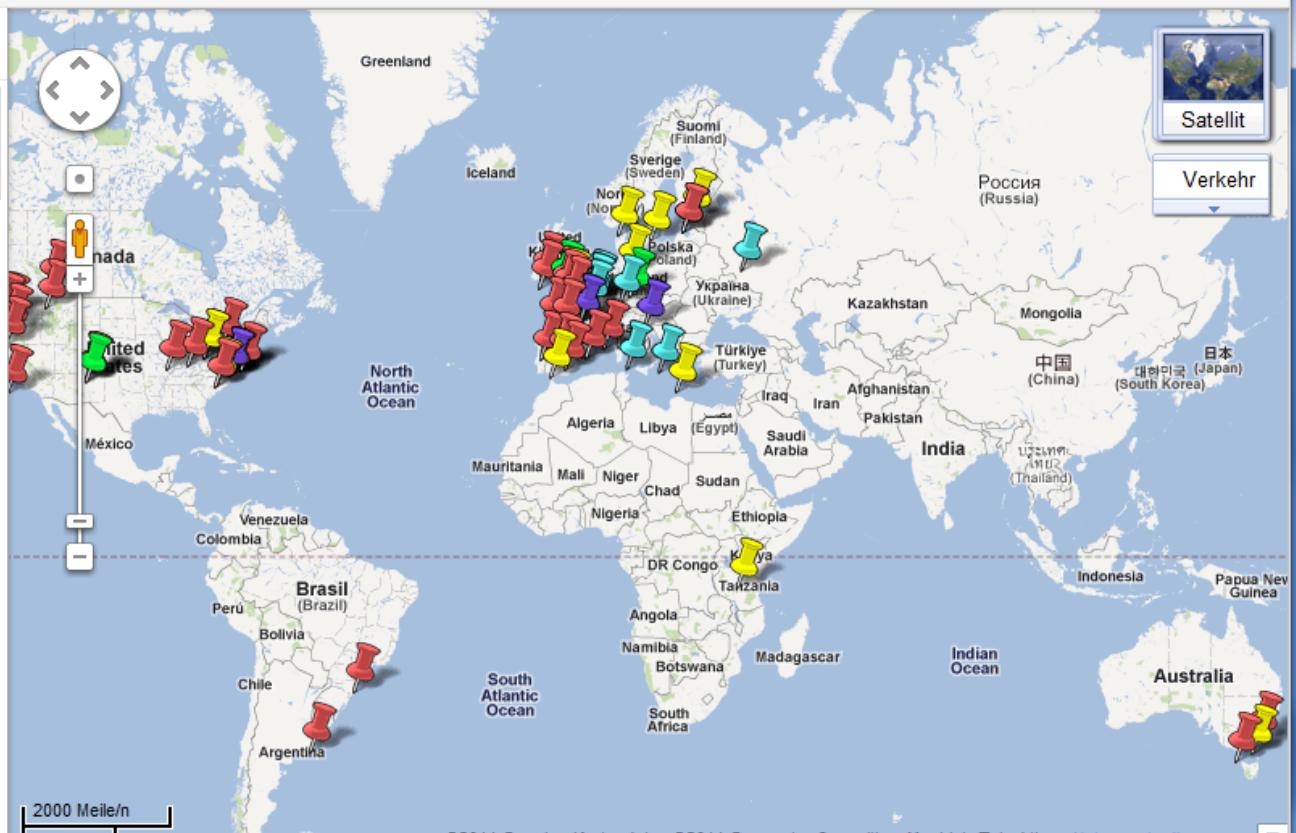
Data are available in a format over which no entity has exclusive control.

8. Licensing

Data are not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed as governed by other statutes.

Finally, compliance must be reviewable.

PLUS: Sunlight Foundation, August 2010: 9) Permanence & 10) Usage Costs



World Map of Open Government Data Initiatives

An open map of those public and private Initiatives which provide Open Government Data or running Challenges.

RED ... local or regional governmental authorities
GREEN ... local or regional private initiatives
YELLOW ... nationwide governmental authorities
BLUE ... nationwide private initiatives
PURPLE ... transnational catalogue

Contact: www.semantic-web.at
Data as KML-File at <http://bit.ly/cHD3Yt>

Öffentlich · 4 Mitbearbeiter · 268.091 Ansichten
Erstellt am 22. Jul. 2010 · Von Semantic Web Company · Aktualisiert
[Karte bewerten](#) · 1 Kommentare · KML · [MapLink](#)

[+1](#) 0

Manchester, UK
Initiative: Open Data Manchester, private

Berlin, D
Initiative: OpenBerlin, private by Open Knowledge Foundation
Deutschland Open Data Network e.V. Liquid Democracy e.V.
Open Everything

Colorado, USA

2000 Meile/n
2000 km

©2011 Google - Kartendaten ©2011 Geocentre Consulting, MapLink, Tele Atlas - [Nutzungsbedingungen](#)

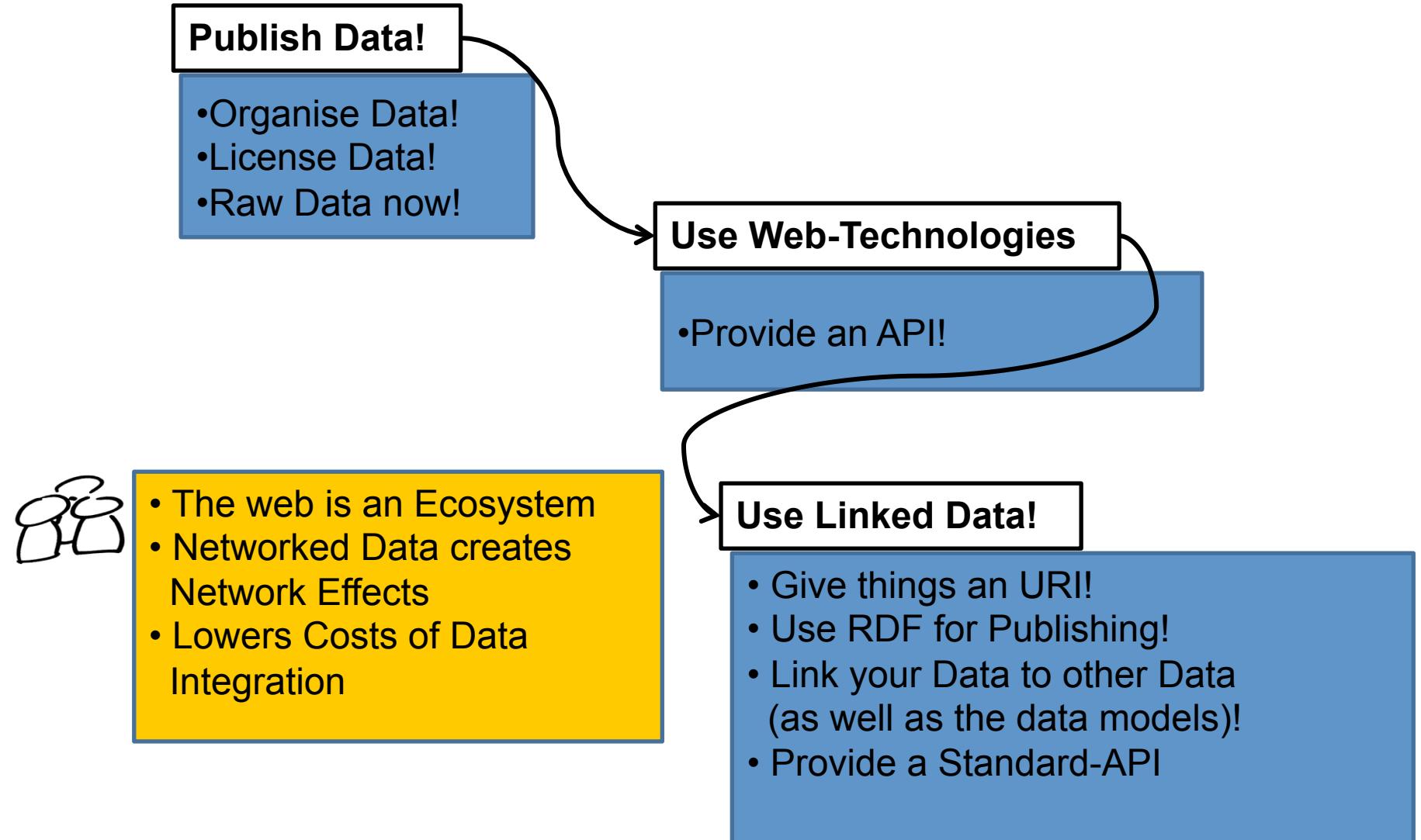
5 Stars for Open Data by Tim Berners Lee

- ★ make your stuff available on the web (whatever format)
- ★★ make it available as structured data (e.g. excel instead of image scan of a table)
- ★★★ non-proprietary format (e.g. csv instead of excel)
- ★★★★ use URLs to identify things, so that people can point at your stuff
- ★★★★★ link your data to other people's data to provide context



©2011 Semantic Web Company -<http://www.semantic-web.at>

7



McKinsey&Company

Report

Big data: The next frontier for innovation, competition, and productivity

May 2011

Data have swept into every industry and business function and are now an important factor of production, alongside labor and capital.

Harnessing big data in the public sector has enormous potential, too. If **US health care** were to use big data creatively and effectively to drive efficiency and quality, the sector could **create more than \$300 billion in value every year**. Two-thirds of that would be in the form of reducing US health care expenditure by about 8 percent.

In the developed economies of **Europe**, **government administrators** could **save more than €100 billion (\$149 billion)** in **operational efficiency improvements** alone by using big data, not including using big data to reduce fraud and errors and boost the collection of tax revenues.

And **users of services enabled by personal location data** could capture **\$600 billion** in consumer surplus.

http://www.mckinsey.com/mgi/publications/big_data/index.asp

EXAMPLES FROM USA / RENSSELAER POLYTECHNIC INSTITUTE

Government Data on the Web



Data.gov - Mozilla Firefox
File Edit View History Bookmarks Tools Help
FRIDAY, AUGUST 07, 2009
<http://www.data.gov/>

DATA.GOV

HOME | CATALOGS | STATE/LOCAL | ABOUT | FAQ | CONTACT US | SUGGEST OTHER DATASETS |

DISCOVER. PARTICIPATE. ENGAGE.

Search the following Data.gov catalogs:

- "RAW" DATA CATALOG
- TOOL CATALOG
- GEODATA CATALOG

FEATURED TOOL:
U.S. GEOLOGICAL SURVEY (USGS)
USGS Global Visualization Viewer for Aerial and Satellite Data

Ten million archive images of the Earth's surface are available for immediate selection and free download via the USGS Earth Resources Observation and Science (EROS) Center's Global Visualization Viewer. Users can preview thumbnails, browse images and download full-image selections from 3.5 million aerial photos of U.S. sites and 8.5 million images captured worldwide by U.S. Earth-observing satellites.

[VIEW THIS TOOL ▶](#)

1 2 3 4 5

Welcome to Data.gov

The purpose of Data.gov is to increase public access to high value, machine readable datasets generated by the Executive Branch of the Federal Government. Although the initial launch of Data.gov provides a limited portion of the rich variety of Federal datasets presently available, we invite you to actively participate in shaping the future of Data.gov by suggesting additional datasets and site enhancements to provide seamless access and use of your Federal data. Visit today with us, but come back often. With your help, Data.gov will continue to grow and change in the weeks, months, and years ahead.

DATA.GOV Data Policy | Accessibility | Contact Info | Privacy Policy

staab@uni-koblenz.de

Recovery.gov - Mozilla Firefox
File Edit View History Bookmarks Tools Help
FRIDAY, AUGUST 07, 2009
<http://www.recovery.gov/>

RECOVERY.GOV

HOME | ABOUT | BOARD | INVESTMENTS | OPPORTUNITIES | IMPACT | NEWS | FAQ | CONTACT US | SIGN UP | SEARCH

CHAIRMAN'S CORNER

Earl E. Devaney is chairman of the Recovery Accountability and Transparency Board. The Recovery Board oversees Recovery Act spending and manages this website. His next report to the American people will focus on development of the new state-of-the-art Recovery.gov.



[LEARN MORE ▶](#)

Data and Statistics – General Reference Resources: USA.gov - Mozilla Firefox
File Edit View History Bookmarks Tools Help
FRIDAY, AUGUST 07, 2009
<http://www.usa.gov/Topics/Refere...>

USA.gov

Home | Site Index | FAQs | E-mail | Phone | Chat | Our Blog | Mobile | Español | Other Languages

Government Web Images News Maps USA.gov

For Citizens | For Businesses and Nonprofits | For Government Employees | For Visitors to the U.S.

Change text size: A A A

E-mail this page | Print this page | Receive updates by e-mail | USA.gov RSS feeds | Tutorials: Find government info | SHARE

By Organization: A-Z Agency Index, Federal Government, State Government, Local Government, Tribal Government

Contact Government: Online Services, Frequently Asked Questions, E-mail, Web Chat, Phone, In Person

2002 Census of Governments, Aging Statistics, American Factfinder, Budget of the U.S. Government, Budget of the U.S. Government: State-by-State Tables FY 2009, Census Data, Child and Family Statistics

Featured Sites: Data.gov

More Reference Materials

Federal IT Dashboard - Mozilla Firefox
File Edit View History Bookmarks Tools Help
FRIDAY, AUGUST 07, 2009
<http://it.usaspending.gov/>

IT DASHBOARD beta

HOME | INVESTMENTS | DATA FEEDS | ANALYSIS | FAQ |

Department of Defense Major Investments: 62 Spending on Major Investments: \$9.6 B (FY 2009)

your window into the federal IT portfolio

DOD DHS HHS DOC Treasury DOT DOJ VA USDA Energy Others

Overall Rating

Agency Investment Spotlight

Government Data Sharing



January 1, 2009
“Openness will strengthen our democracy and promote efficiency and effectiveness in Government.”
--- President Obama



May 21, 2009
data.gov online
57 Data Sets



December 8, 2009
“Open Government Directive” released
~2000 Data Sets



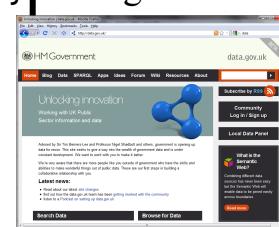
May 21, 2010
data.gov relaunch with semantic web featured
>305,000 Data Sets

2010 ...

~6000 Data Set



June 30, 2009
Putting Govt Data online-
Data.gov.uk beta



January 19, 2010
data.gov.uk online

Important to the citizens: eg. Education



You're logged in as [jaggeroe](#) (logout)

[Create a pipe](#)

data.gov.uk Education data demo

Pipe Web Address: <http://pipes.yahoo.com/useful/baae2049ef8831992c657d6cf486bec>

[View Source](#) [Clone](#)

Query `SELECT ?school ?name ?` [Run Pipe](#)

Configure this Pipe

Get as a Badge [MY YAHOO!](#) [Google](#) [Get as RSS](#) [Get as JSON](#) [More options](#)

[Map](#) [List](#)

10 items

Forge Valley Community School
Opening date: Closure date:

Properties
Not published 0 runs
7 clones

Bookmark / Share
[F](#) [G](#) [T](#) [DZ](#) [C](#)

Modules (8)
fetchdata
regex
rename
textinput
strconcat
subpipe
loop
yql

[View Source](#)

Map Sat Hyb NAVTEQ Data ©2009 Yahoo! Inc.

Report abusive Pipe this is YAHOO!

Your Privacy | Terms of Use | Copyright/IP Policy | Help | Feedback
Copyright © 2009 Yahoo! Inc. All rights reserved.

Data.gov.uk

New York State School Report Cards

[View](#) [Edit](#) [Revisions](#) [Track](#)

Description: This maps each public school in the state and offers a look into some of the statistics available. This is more of a proof of concept, as the data is somewhat limited.

Population: 294,565 (2000)
Date founded: 1683
Household Count: 120,645 (2008)
BLS Unemployment Rate (%): 0.0533 (2008)
Latitude: 42.6621
Longitude: -73.8491
Source: http://www.firebaseio.com/view/en/albany_county

School Info

Demographic Information by Percentage

Year	Ameri.	Black	Hispa.	Asian	White	Multi...
2007	~75%	~10%	~5%	~5%	~5%	~5%
2008	~75%	~10%	~5%	~5%	~5%	~5%
2009	~75%	~10%	~5%	~5%	~5%	~5%

High School Completers by Percentage

Year	Percentage
2007	~85%
2008	~85%
2009	~85%

Contributor: [webscience_spring2010_5](#)

Map Data ©2009 Google Map Data ©2009 Google

RPI NYS demos

Moving data.gov to linked data (UK)



- Built around “linked data” from the start
- Authorization for this from the Prime Minister

The screenshot shows the homepage of the hmg.gov.uk/data website. At the top, there's a navigation bar with links for Home, Data, Blog, Apps, Ideas, Resources, and About. The Home link is highlighted in orange. To the right of the navigation is a search bar with the placeholder "Search data.gov.uk" and a "Subscribe by RSS" button with an orange icon. Below the navigation, the main content area features a large blue 3D molecular or atom-like graphic. The title "Unlocking innovation" is displayed above the text "Working with UK Public Sector information and data". A paragraph explains that the site is advised by Sir Tim Berners-Lee and Professor Nigel Shadbolt, and aims to open up government data for reuse. Another paragraph expresses the desire to work with others to make it better and build a collaborative relationship. A "Search Data" section includes a search input field and a "Search" button. A "Featured Data" section highlights "Education and training participation: 16-18 year olds" with a brief description and a "View all apps" link. A "Most Recent Apps" section lists three apps: "PlanningAlerts" (PlanningAlerts.com is a...), "FillThatHole" (This site provides a single...), and "Where Can I Live?" (An app to tell you where...). On the right side, there's a "Digital Engagement Twitter stream" showing a tweet from RT @innovate_uk about Innovate 09, a "Submit an app" section with a "Submit your app" button, and a "Featured blogs" section.

Moving data.gov to linked data (US)



- Third parties (like RPI) translate the government datasets into linked data formats

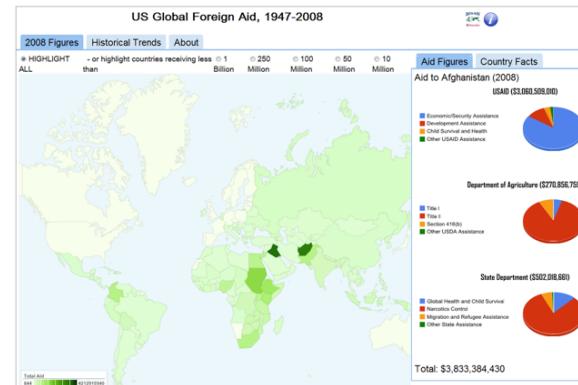
The screenshot shows the homepage of Data.gov. At the top, there's a navigation bar with links for HOME, DATA, APPS, COMMUNITY, METRICS, OPEN DATA SITES, GALLERY, and WHAT'S NEW. A search bar is also present. The main content area features a large image of two \$100 bills. Below this, there's a section titled "SBA SMALL BUSINESS LOANS & GRANT PROGRAM" with a brief description and a link to the dataset. To the right, a sidebar lists "Latest Datasets" including U.S. Trade Overview Presentation, U.S.-Sub-Saharan Trade Report/Profile, Renewable Energy and Energy Efficiency, Online list of foreign safeguard measures, Online FTZ Information System (OFIS), KORUS Industry-by-Industry Fact Sheet, FTZ Manufacturing (T/M) Database, July 2011 USCIS FOIA LOG, June 2011 USCIS FOIA LOG, and May 2011 USCIS FOIA LOG. The page is divided into several sections: "DATA AND APPS" (with a thumbnail of a map and a list of datasets), "COMMUNITIES" (with a world map and a link to the Semantic Web community), "OPEN GOVERNMENT" (with a thumbnail of the American flag and news about Japanese Earthquake and Radiation Data), "LEARN" (with a thumbnail of children at a computer and a link to classroom resources), "SEMANTIC WEB" (with a thumbnail of a molecular structure and a link to the Semantic Web community), and "DEVELOPER'S CORNER" (with a thumbnail of mobile devices and a link to developer resources). Each section has a "MORE" button.

- US Data.gov hosts 6.4B RDF triples 5/21/2010
 - Semantic Web community hosted
 - <http://data.gov/semantic>

Linked data lets us create “Data” Mashups



(a) White House visitor search



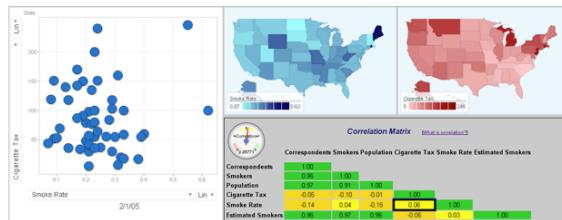
(b) US-UK Foreign Aid Comparison



(c) Agency Budget and NYTimes



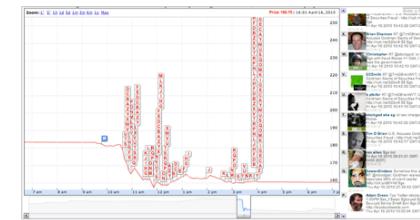
(d) Wildland fire and DBpedia



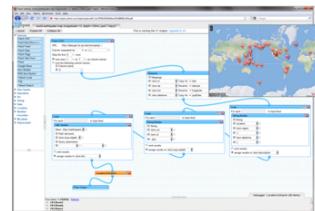
(e) [Health] Tobacco Prevalence and Correlated Factors



(f) [Policy] About Supreme Court Justices



(g) [Financial] Stock price and Twitter events



(h) [Yahoo! Pipes] World Earthquake Map



(i) [IBM ManyEyes] White House visitor network

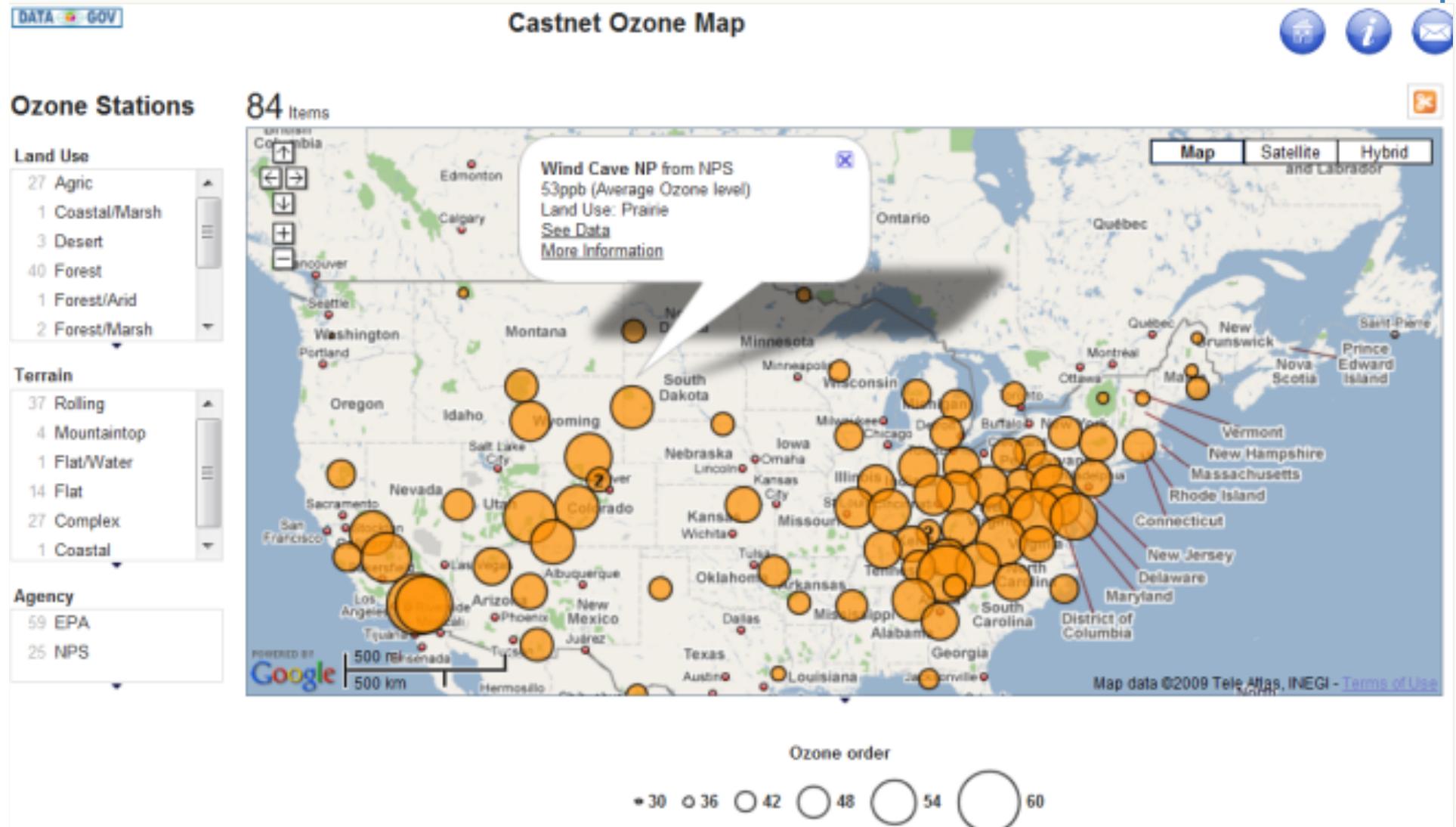


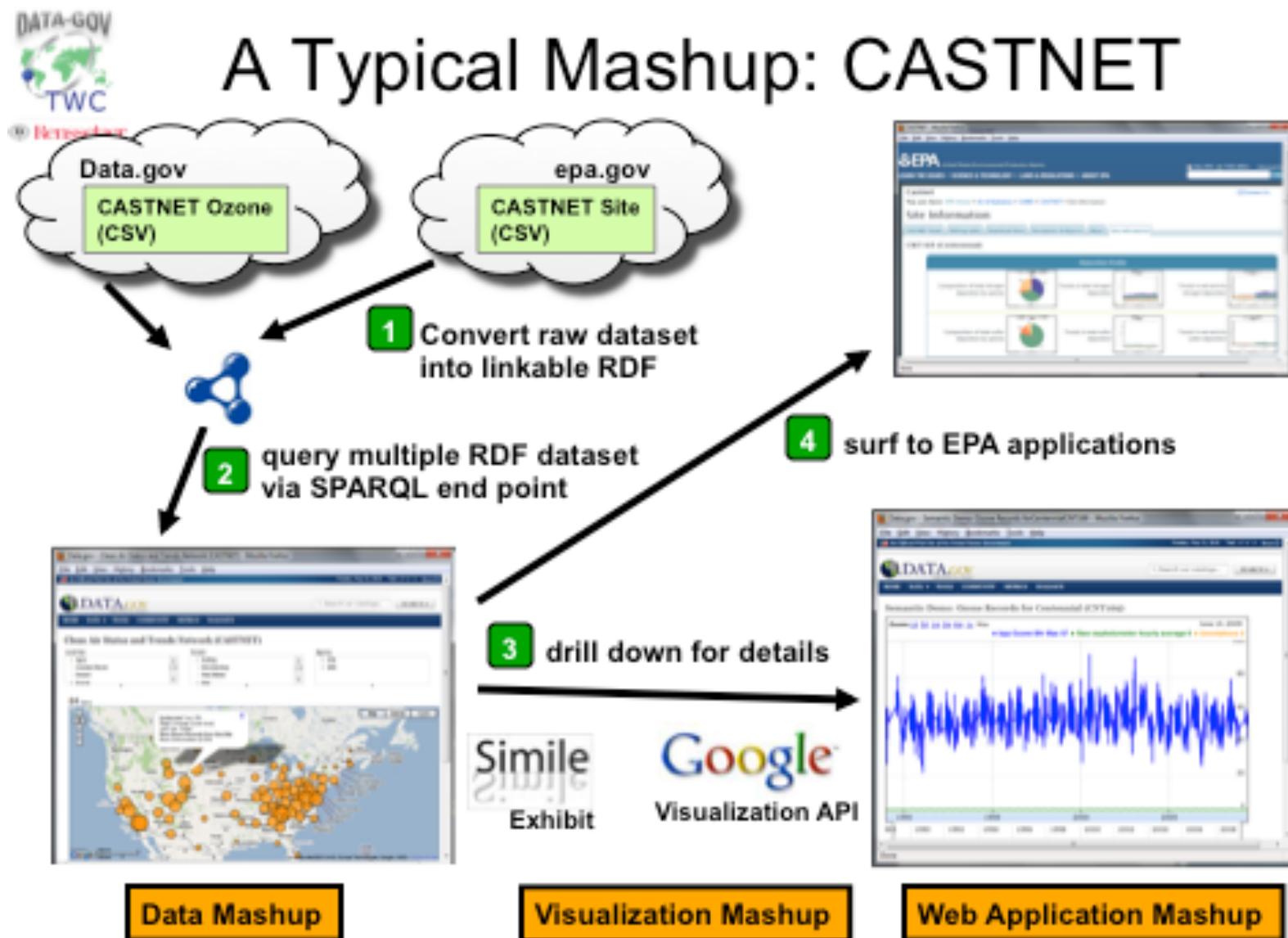
(j) [RDFa] semantic search



(k) [RSS] data.gov updates

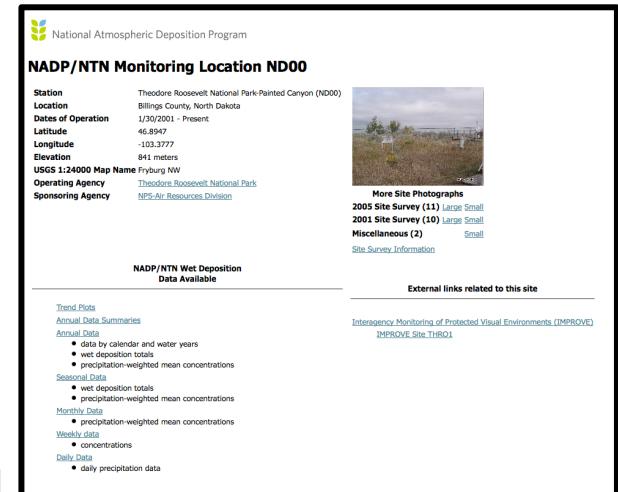
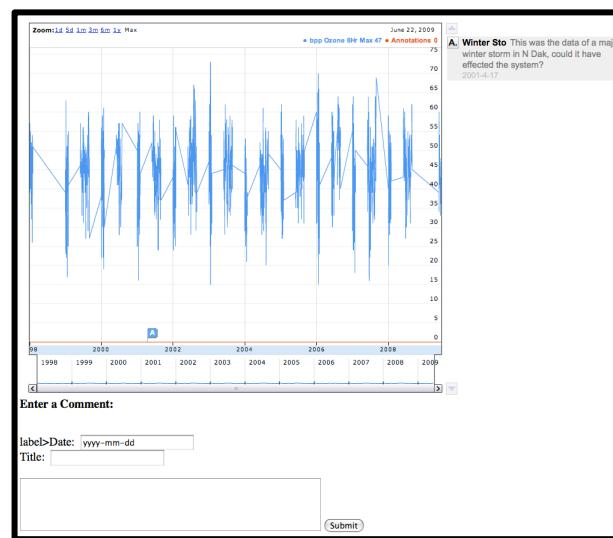
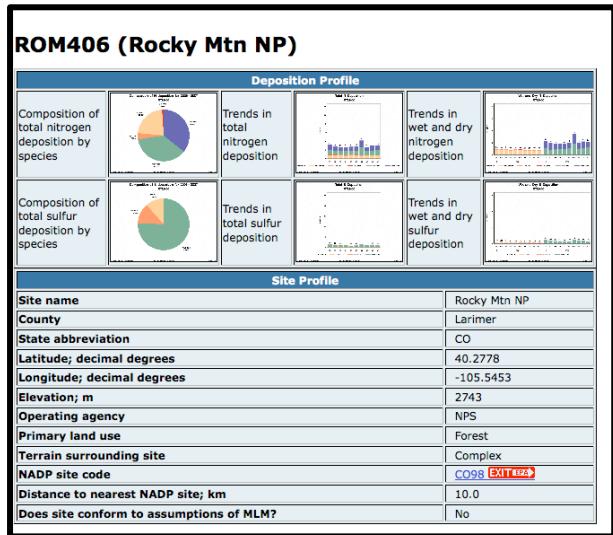
More than 50 of these at <http://logd.tw.rpi.edu>
(and lots more at data.gov.uk)





Created by Dominic DiFranzo, PhD student at RPI, <http://www.data.gov/semantic/Castnet/html/exhibit>

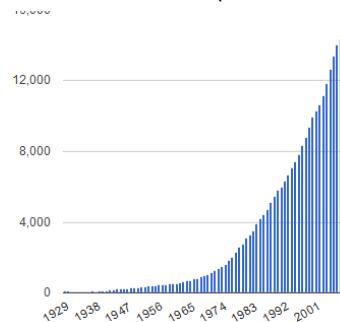
Adding some Web magic



Linking GDP of the US and China

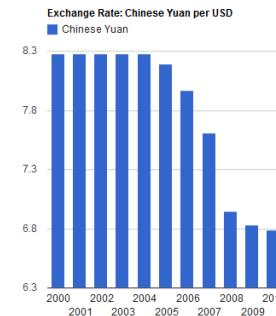


GDP of the US (Billion Dollar)



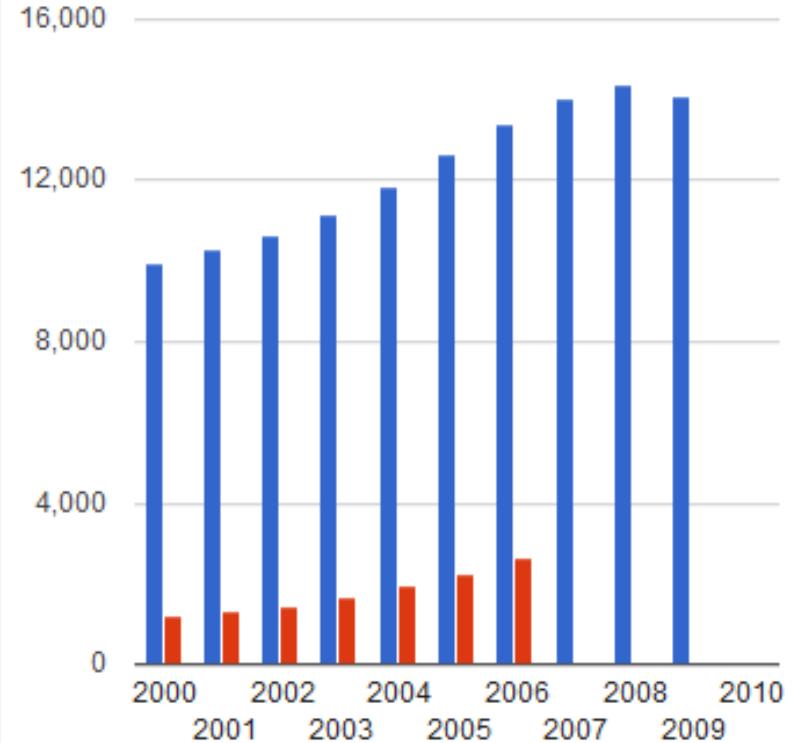
Federal Reserve Statistical Release

H.10
Foreign Exchange Rates (Weekly)

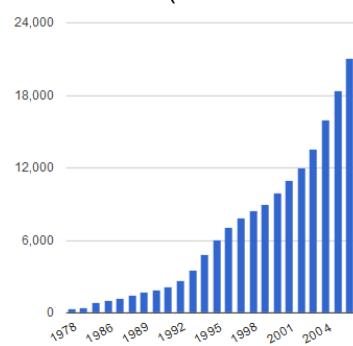


GDP (Billion USD) adjusted by Exchange Rate

US China



GDP of China (Billion Chinese Yuan)



[Temporal Mashup] bea.gov + federalreserve.gov + stats.gov.cn

Linking GDP of the US and China

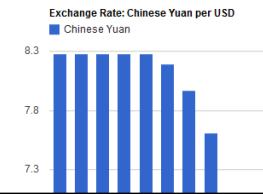


GDP of the US (Billion Dollar)



Federal Reserve Statistical Release

H.10
Foreign Exchange Rates (Weekly)



GDP (Billion USD) adjusted by Exchange Rate

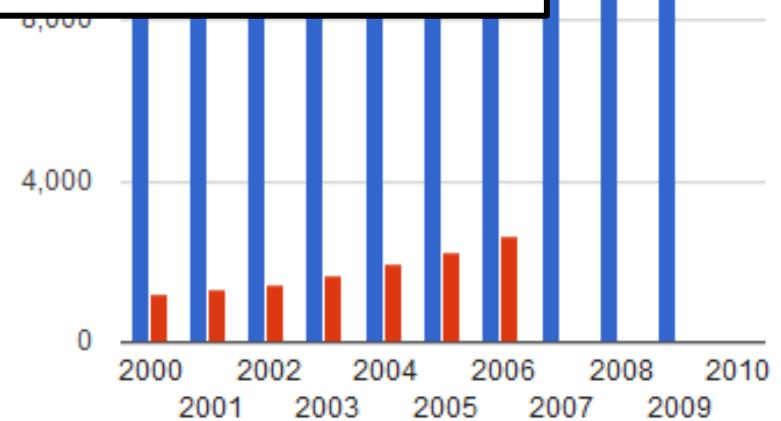
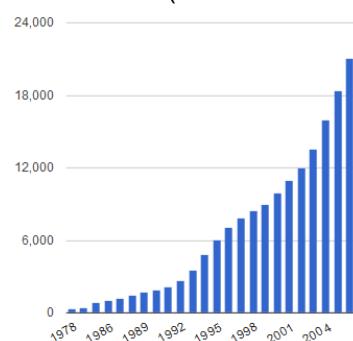
US China

16,000

This mashup was built in less than 4 hours – including conversion of data, web interface, and visualization!



GDP of China (Billion Chinese Yuan)

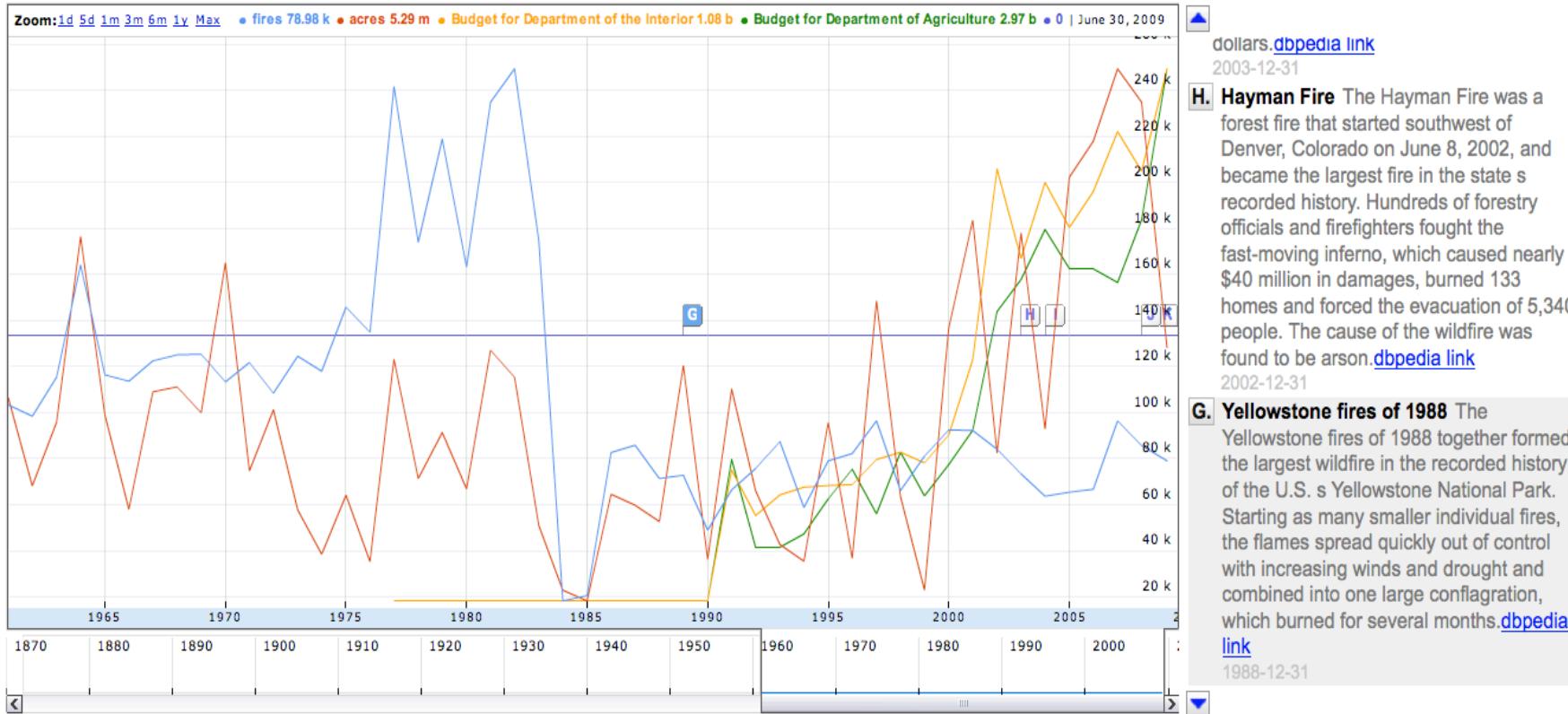


[Temporal Mashup] bea.gov + federalreserve.gov + stats.gov.cn

Govt systems can use linked data web •for context



National Wild Fire Statistics and Budget



Datasets: acres burned, and agency budgets
Dbpedia: wikipedia descriptions of major US fires

Integrate with Social media



Linking Open Government Data

Home Team Endpoint View Edit Revisions Track View Source

Description: This map shows some locations that receive funding from the New York State Government. Click on a pin to learn more about that location and cast your votes via twitter.

Cast Your Votes on New York State Projects

238 Items

40 results out of 238 cannot be plotted.

Map Satellite Hybrid

Search Projects

City

- 1 Albertson
- 2 Baldwin
- 6 Bellmore
- 7 Bethpage
- 1 East Hills

Agency

- 16 Aging
- 6 ASAS
- 6 C&F
- 53 CFS
- 1 DCJS

East Meadow LAX

Description: To help defray operating expenses and provide programs for the youth.

Address: 2510 Second Avenue
East Meadow, NY 11554

Funding: \$1750

Senator: Hannon

For: [Tweet](#) 2

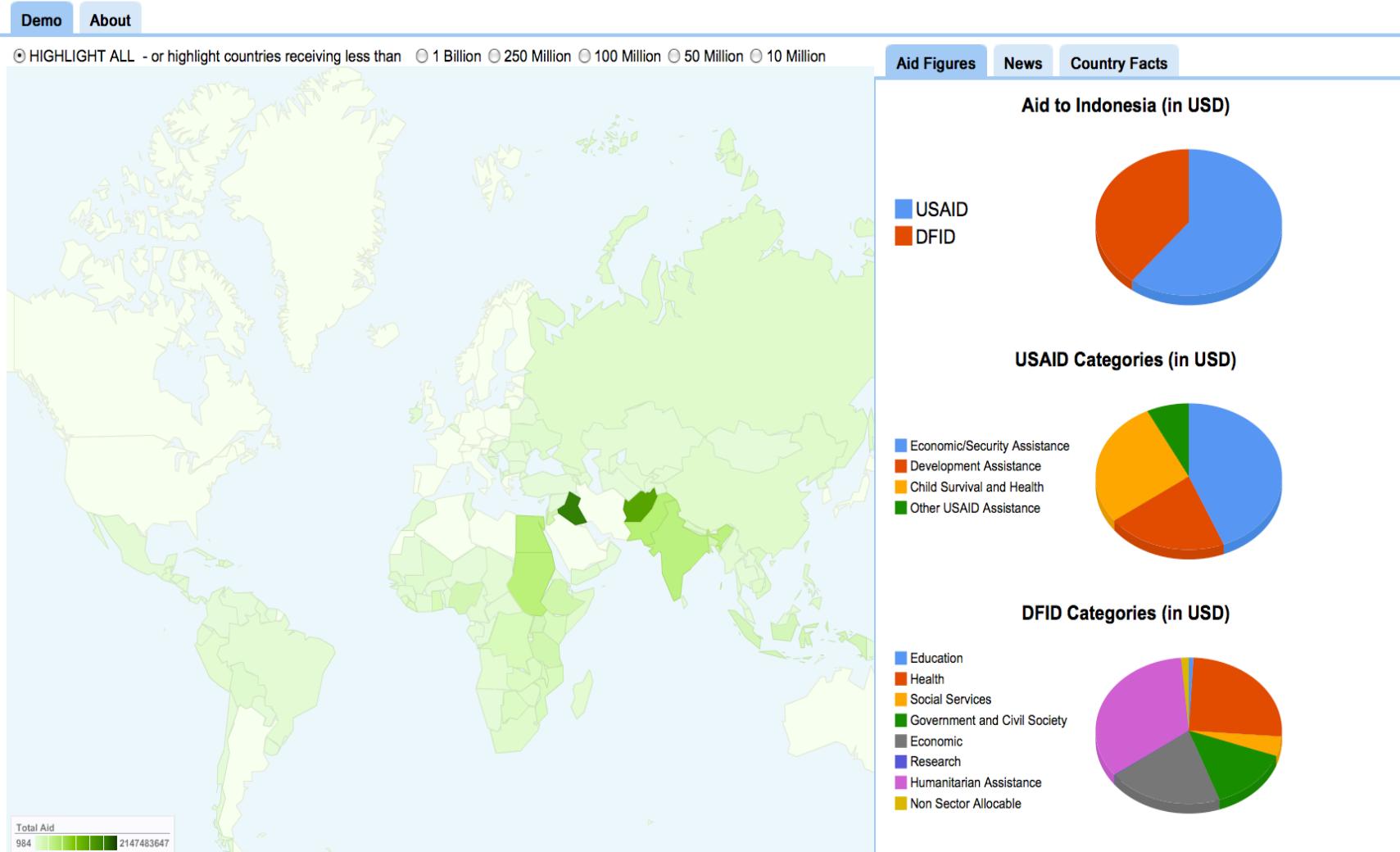
Against: [Tweet](#) 2

A map of Long Island, New York, showing various locations marked with orange pins. A callout box provides detailed information about a project in East Meadow, LAX. The map includes labels for Jericho Turnpike, Northern State Pkwy, Old Country Rd, Corporate Dr, Merrick Rd, Stewart Ave, Meadowbrook State Pkwy, Hempstead Plains Preserve, and Levittown. Roads are labeled with numbers like 25, 24, and 107.

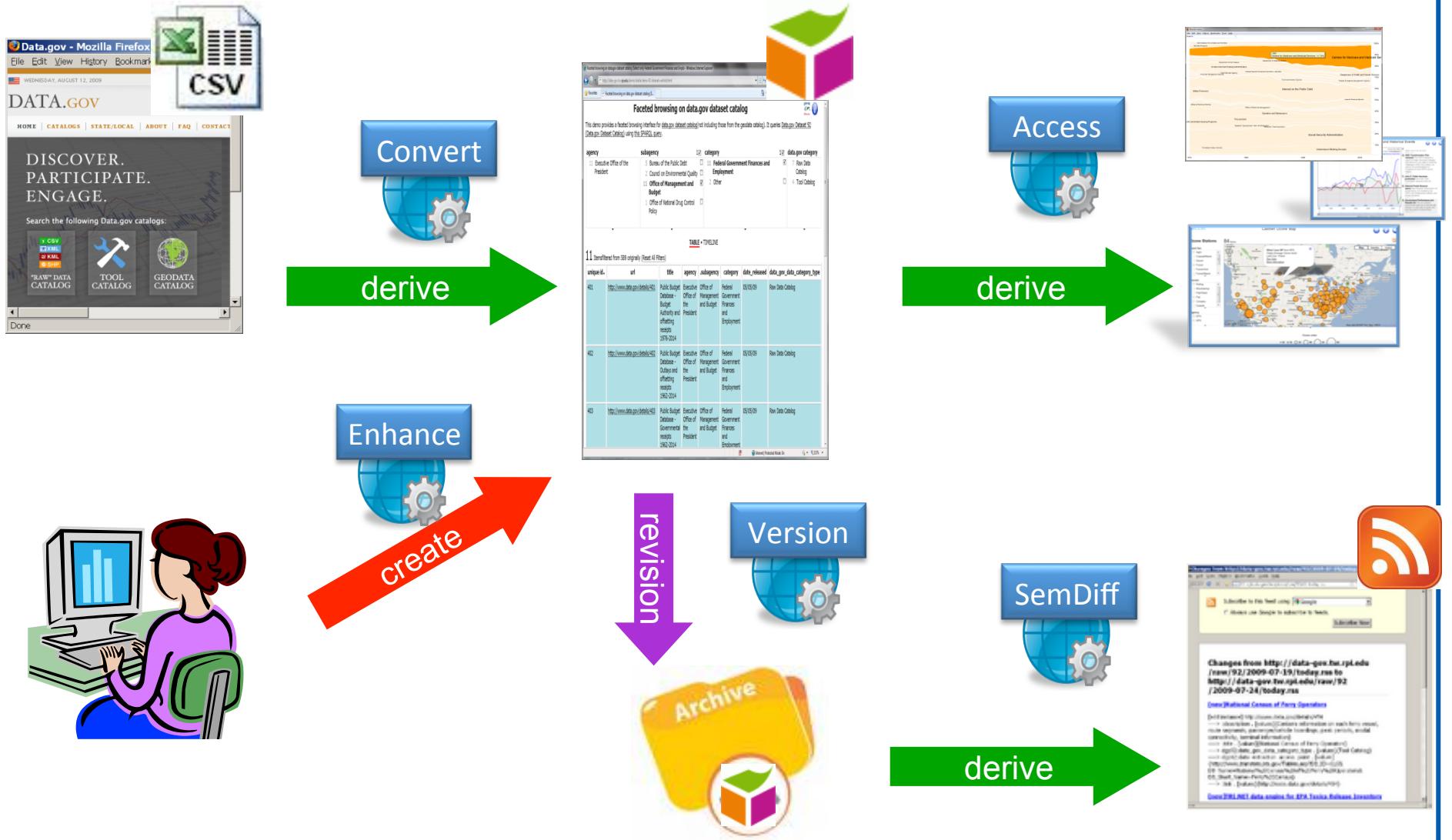
Combining data from different data sharing sites



Comparing US (USAID) and UK (DFID) Global Foreign Aid



RPI workflow enhances raw RDF w/ useful URIs



 Rensselaer

Linking Open Government Data

Menu

- Home
- International Dataset Search
- Instance Hub
- SW Challenge 2010
- Tutorials
- Videos
- Demos
- Datasets
- Tools & Technologies
- Community
- About Us
- Publications
- RSS Feed

User login

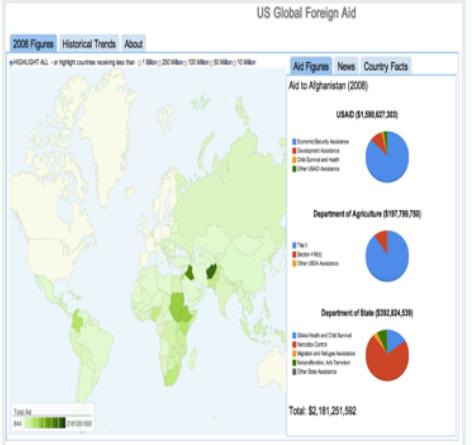
Username: *

Password: *

Linking Open Government Data

[View](#) [View Source](#)

Featured LOGD Demos



Aid Figures | Historical Trends | About

US Global Foreign Aid

HIGHLIGHT ALL → highlight countries receiving less than: <1 Miln > 200 Miln > 100 Miln > 50 Miln > 10 Miln

Aid Figures News Country Facts

Aid to Afghanistan (2008)

- USAID (\$1,086,871,381)
- Economic Recovery Assistance
- Development Assistance
- Child Survival and Health
- Other USAID Assistance

Department of Agriculture (\$197,796,750)

- Farm Bill
- Food for Peace
- Other USDA Assistance

Department of State (\$212,824,453)

- Global Health and Child Survival
- Migrant and Refugee Assistance
- Other Development Assistance
- Other State Assistance

Total: \$2,181,291,582

Contributor(s):



US Global Foreign Aid

LOGD Portal Updates (aggregated RSS Feeds from several sources)

LOGD Videos (source: LOGD site)

Analyzing Political Topics Using Open Government Data
by dingl - Sep 27, 2010

Quick Links

- Visit our new [International Dataset Catalog](#)
- **Semantic Web Challenge 2010 - TWC wins 2nd prize at SWC 2010!!!**
- [Demos](#) - Examples of consuming open government data
- [Tutorials](#) - Learn Semantic Web technologies for LOGD
- [Feedback](#) - Send us comments
- [Issue tracker](#) - Report a bug
- [LOGD Open Source on Google Code](#)
- [LOGD SPARQL Endpoint \(Alpha\)](#)

TWC LOGD Datasets Stats (as of 2011-11-04)

- TWC LOGD is hosting **9,951,771,397** RDF triples
- We have created **1,888** RDFized datasets originating from **119** sources.
- We have enhanced **1,686** RDFized datasets using **9,499** links to other LOD datasets.
- For more statistics, see [TWC LOGD: Data Description and Statistics](#).

Linking Open Government Data

Rensselaer

TWC LOGD Datasets Stats (as of 2011-11-04)

- TWC LOGD is hosting **9,951,771,397** RDF triples
- We have created **1,888** RDFized datasets originating from **119** sources.
- We have enhanced **1,686** RDFized datasets using **9,499** links to other LOD datasets.
- For more statistics, see [TWC LOGD: Data Description and Statistics](#).

View International Dataset Catalog
Web Challenge 2010 - TWC wins 2nd prize
2010!!!
Examples of consuming open government data
Learn Semantic Web technologies for LOGD
Send us comments
Poster - Report a bug
Open Source on Google Code
SPARQL Endpoint (Alpha)

Datasets Stats
(as of 2011-11-04)

is hosting **9,951,771,397** RDF triples

- We have created **1,888** RDFized datasets originating from **119** sources.
- We have enhanced **1,686** RDFized datasets using **9,499** links to other LOD datasets.
- For more statistics, see [TWC LOGD: Data Description and Statistics](#).

Username: *
Hendler

Password: *

Log in

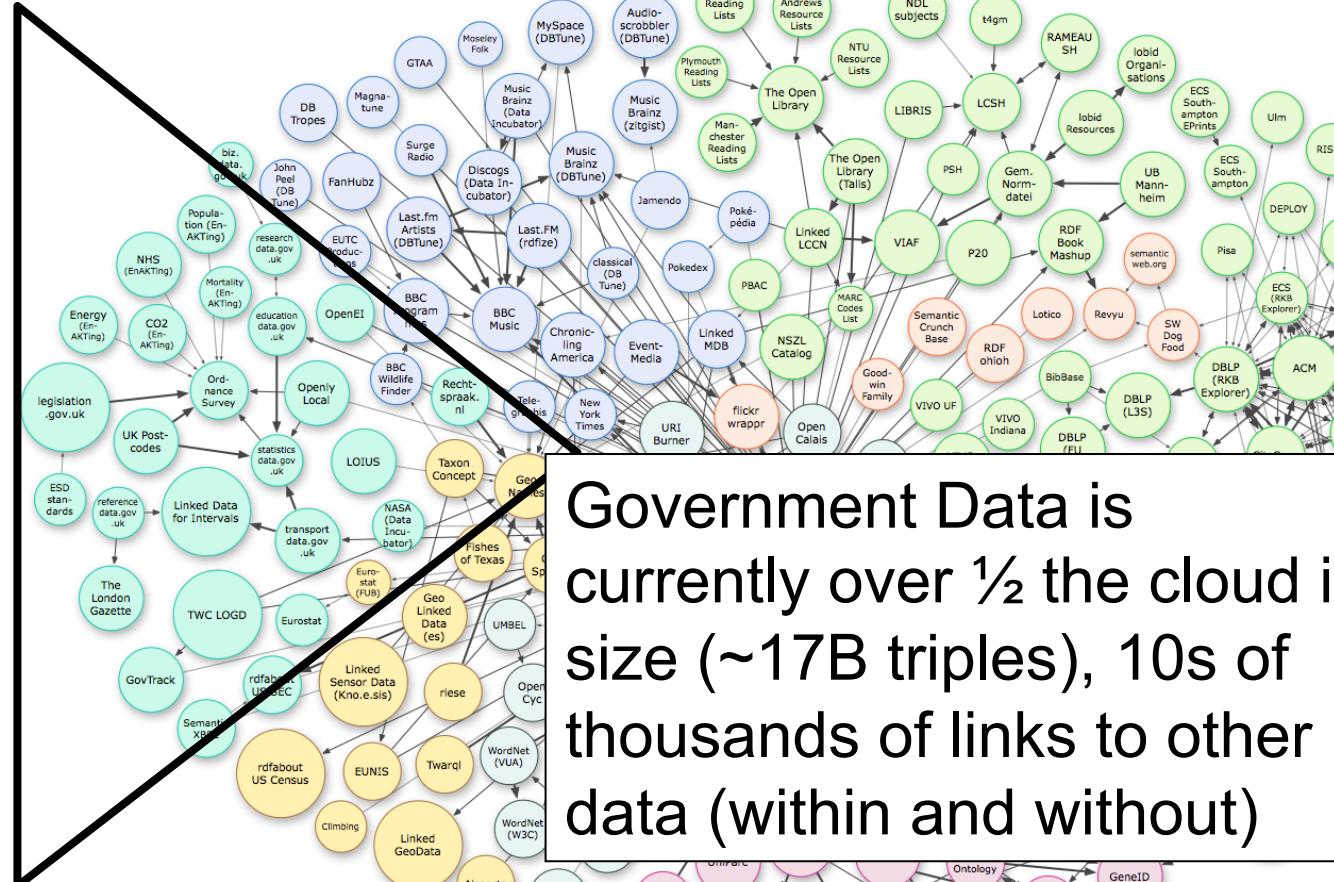
US Global Foreign Aid

LOGD Portal Updates (aggregated RSS Feeds from several sources)

LOGD Videos (source: LOGD site)

Analyzing Political Topics Using Open Government Data
by dingl - Sep 27, 2010

Government Data in the linked open data cloud



As of September 2010

- URI design is crucial to govt data sharing
 - ◆ esp. within govts
 - ◆ Whether your goal is linked data or not
- UK Government has designed and made great use of standard URI practices in their linked data
 - ◆ US exploring URI design schemes
 - Join the community at Semantic.data.gov and participate!

TWC LODG Instance Hub

[View](#)[View Source](#)

A well-known benefit of using [Linked Data](#) principles when publishing government data is the ease with which the resulting data may be integrated and linked with diverse other datasets. Realizing such benefits ultimately hinges on unambiguously identifying *common entities* within these datasets. From a practical standpoint this means identifying which literals represent instances of entities --- US states, US government agencies, toxic chemicals, crops, etc --- and associating *canonical URIs* with those occurrences across datasets. To facilitate interlinking, it is then helpful to relate these literals and our newly-minted canonical identifiers to other identifiers in common usage --- DBpedia data pages, information pages used in various industries, etc --- to further facilitate interlinking with other datasets.

The **TWC LODG Instance Hub Project** seeks to demonstrate the utility of this approach for entities occurring across diverse government datasets. First, we have identified several categories of entities from Data.gov datasets. We have designed canonical URIs based on a draft set of [principles](#) that may be applied across a wide range of government data sources (agencies, etc). And we have provided links to synonymous identifiers for these entities in common use in other systems to facilitate interlinking.

TWC has established a draft set of [URI Design Principles](#) that government agencies may choose to adopt as they establish authoritative, canonical URIs for entities of special concern within their datasets. The TWC LODG Instance Hub demonstrates how to associate these canonical URIs with instances of entities that occur as literals or as URIs from other naming systems.

The following table provides links to instance data for a number of categories derived from US Government data via [Data.gov](#). Click on a category to browse these instances or download the RDF data (in Turtle).

Instance Hub Category	RDF Dump
US States	.ttl
US Agencies	.ttl
US Fiscal Years	.ttl
Crops	.ttl
Toxic Chemicals	.ttl
Alternative Transport Fuels, etc (Coming Soon)	

Example: US States

New York

Instance Hub: New York

http://lodg.tw.rpi.edu/id/us/state/New_York

[Instance Hub State Page](#)

[Instance Hub Main Page](#)



http://dbpedia.org/resource/New_York

<http://sw.opencyc.org/concept/Mx4rvViNs5wpEbGdrcN5Y29ycA>

Linked To: <http://rdf.freebase.com/ns/m/059rby>

<http://rdf.freebase.com/ns/m/01mf5p>

<http://www.rdfabout.com/rdf/usgov/geo/us/ny>

Datasets Linked:

http://lodg.tw.rpi.edu/source/eia-doe-gov/dataset/fossil_fuel_consumption/version/2010-Jul-20

http://lodg.tw.rpi.edu/source/census-gov/dataset/us_regdiv/version/2000-Sep-07

<http://lodg.tw.rpi.edu/source/data-gov/dataset/1292/version/2009-May-18>

<http://lodg.tw.rpi.edu/source/data-gov/dataset/1149/version/1st-anniversary>

<http://lodg.tw.rpi.edu/source/data-gov/dataset/1148/version/1st-anniversary>

<http://lodg.tw.rpi.edu/source/data-gov/dataset/1623/version/2009-May-18>

<http://lodg.tw.rpi.edu/source/data-gov/dataset/92/version/2011-Jan-30>

<http://lodg.tw.rpi.edu/source/twc-rpi-edu/dataset/instance-hub-us-states-and-territories/version/2011-Mar-31>

<http://lodg.tw.rpi.edu/source/epa-gov/dataset/national-air-toxics-trends-stations/version/2010-Sep-24>

<http://load.tw.rpi.edu/source/data-qov/dataset/4383/version/2011-Nov-29>

Example: US Govt Agencies



National Oceanic and Atmospheric Administration

Instance Hub: National Oceanic and Atmospheric Administration

http://lod.tw.rpi.edu/id/us/fed/agency/Department_of_Commerce/National_Oceanic_and_Atmospheric_Administration

Instance Hub US Agency Page

Instance Hub Main Page

Abbreviation: NOAA; DOC/NOAA

The National Oceanic and Atmospheric Administration (NOAA), pronounced /noʊə.ə/, like "noah", is a scientific agency within the United States Department of Commerce focused on the conditions of the oceans and the atmosphere. NOAA warns of dangerous weather, charts seas and skies, guides the use and protection of ocean and coastal resources, and conducts research to improve understanding and stewardship of the environment. In addition to its civilian employees, NOAA research and operations are supported by 300 uniformed service members who make up the NOAA Commissioned Officer Corps. The current Under Secretary of Commerce for Oceans and Atmosphere at the Department of Commerce, and the agency's Administrator, is Dr. Jane Lubchenco, nominated by President Barack Obama and confirmed by the United States Senate on March 19, 2009.

Homepage: <http://www.noaa.gov/>

owl:SameAs: http://dbpedia.org/resource/National_Oceanic_and_Atmospheric_Administration
<http://rdf.freebase.com/ns/m/09fyq>



Datasets Linked:

- <http://lod.tw.rpi.edu/source/data-gov/dataset/92/version/2011-Jan-30>
- <http://lod.tw.rpi.edu/source/twc-rpi-edu/dataset/instance-hub-us-federal-agencies/version/2011-Jul-7>
- <http://lod.tw.rpi.edu/source/data-gov/dataset/92/version/2011-Jul-28>
- http://lod.tw.rpi.edu/source/twc-rpi-edu/dataset/instance-hub-us-federal-agencies/version/2011-Mar-31_17-04-26
- <http://lod.tw.rpi.edu/source/data-gov/dataset/92/version/2010-Dec-11>
- <http://lod.tw.rpi.edu/source/data-gov/dataset/92/version/2011-Jun-13>



Instance Hub: soybean

<http://lod.tw.rpi.edu/source/usda-gov/id/us/crops/soybean>

[Instance Hub US Crop Page](#)

[Instance Hub Main Page](#)

Crop Category:	oil crops
USDA Briefing URL:	http://www.ers.usda.gov/Briefing/SoybeansOilCrops/ http://dbpedia.org/resource/Soybean
owl:SameAs:	http://eunis.eea.europa.eu/species/171391 http://rdf.freebase.com/ns/m/0gyv9 http://lod.geospecies.org/ses/52Dxq

Datasets Linked:

<http://lod.tw.rpi.edu/source/twc-rpi-edu/dataset/instance-hub-us-crops/version/2011-Aug-20>

- Metadata design is crucial to govt data sharing
 - ◆ Needed for search and federation in large data sharing efforts
- International data sharing will be a crucial next step
 - ◆ W3C Govt Linked Data Working Group
 - ◆ Need for vocabularies within govt sectors
 - Esp for cross-language use

International Open Government Data Search



International OGD Catalog Search, searching **440,386** datasets from **116 catalogs** in **16 languages** representing **38 countries** and international organizations

Results 1 to 200 of 440386

Dataset Title	Agency	Categories	Catalog	Country
Elektronischer Wasserstraßen-Informationsservice - [Karte Nordwest]	Niederschsischer Landesbetrieb für Wasserwirtschaft Ksten- und Naturschutz	Wasser	Portalu.de (Germany)	Germany
Hochwasserrückhaltebecken online	Niederschsischer Landesbetrieb für Wasserwirtschaft Ksten- und Naturschutz	Wasser	Portalu.de (Germany)	Germany
Bergbau - Kontrollergebnisse	Landesamt für Geologie und Bergwesen Sachsen-Anhalt LAGB	Weitere	Portalu.de (Germany)	Germany
Lufthygienische Überwachung Schleswig Holstein	Lufthygienische Überwachung Schleswig-Holstein	Luft	Portalu.de (Germany)	Germany
Überwachung der Umweltradioaktivität (IMIS)	Bundesamt für Strahlenschutz	Strahlung	Portalu.de (Germany)	Germany
Hochwasserinfos für Brandenburg	Ministerium für Umwelt Gesundheit und Verbraucherschutz des Landes Brandenburg	Wasser	Portalu.de (Germany)	Germany
Aktuelle Pegel der Elbe in Brandenburg	Ministerium für Umwelt Gesundheit und Verbraucherschutz des Landes Brandenburg	Wasser	Portalu.de (Germany)	Germany
		chemikalien, gesundheit		

next

Catalogs:

- (395459) Data.gov --- Geodata Catalog
- (7702) Data.gov.uk (United Kingdom)
- (5978) Data.gov.sg (Singapore)
- (3958) Opendatacordoba.com(Cordoba,Spain)
- (3707) Data.gov (United States) --- Raw Data
- (2489) Data.gov.bc.ca (British Columbia, Canada)
- (1616) Data.govt.nz (New Zealand)
- (1165) Data.medicare.gov (USA)

Countries:

- (406280) United States
- (8940) United Kingdom
- (5978) Singapore
- (5907) Spain
- (4777) Canada
- (1687) New Zealand
- (1293) Italy
- (776) Australia

Agencies:

- (2271) Ayuntamiento de Cordoba
- (1666) Environmental Protection Agency
- (1468) Health
- (1347) Economic Development Board
- (1123) Department of Health
- (1117) U.S. Geological Survey Earth Resource
- (883) Land Information New Zealand
- (781) Department for Communities and Local Government
- (758) Regione Autonoma della Sardegna

Categories:

- (147) Energy and Transport
- (140) Datos Demográficos
- (139) Empleo
- (139) Infraestructuras
- (136) Labour
- (134) Finance
- (129) Sciences
- (126) Transportation and Storage
- (126) State sector performance

There's lots of data out there!!

International OGD Catalog Search, searching 440,386 datasets from 116 catalogs in 16 languages representing 38 countries and international organizations

Results 1 to 200 of 440386

Dataset Title	Agency	Categories	Catalog	Country
Elektronischer Wasserstraßen-Informationsservice - [Karte Nordwest]	Nieders chsischer Landesbetrieb f r Wasserwirtschaft K sten- und Naturschutz	Wasser	Portalu.de (Germany)	Germany

International OGD Catalog Search, searching 440,386 datasets from 116 catalogs in 16 languages representing 38 countries and international organizations

- Catalogs:
- (395459) Data.gov --- Geodata Catalog
 - (7702) Data.gov.uk (United Kingdom)
 - (5978) Data.gov.sg (Singapore)
 - (3958) Opendatacordoba.com(Cordoba,Spain)
 - (3707) Data.gov (United States) --- Raw Data
 - (2489) Data.gov.bc.ca (British Columbia, Canada)
 - (1616) Data.govt.nz (New Zealand)
 - (1155) Data.medicago.gov (USA)

- Agencies:
- (4777) Canada
 - (1687) New Zealand
 - (1293) Italy
 - (776) Australia

- Categories:
- (2271) Ayuntamiento de Cordoba
 - (1666) Environmental Protection Agency
 - (1468) Health
 - (1347) Economic Development Board
 - (1123) Department of Health
 - (1117) U.S. Geological Survey Earth Resource
 - (883) Land Information New Zealand
 - (781) Department for Communities and Local Government
 - (758) Regione Autonoma della Sardegna

- Chemical, environmental and health:
- (140) Datos Demográficos
 - (139) Empleo
 - (139) Infraestructuras
 - (136) Labour
 - (134) Finance
 - (129) Sciences
 - (126) Transportation and Storage
 - (126) State sector performance

- Faceted browser with
 - ◆ Keyword search
 - ◆ Catalogs
 - ◆ Countries
 - ◆ Agencies
 - ◆ Categories
 - ◆ (in any order)

International OGD Catalog Search, searching 440,386 datasets from 116 catalogs in 16 languages representing 38 countries and international organizations

Results 1 to 200 of 440386 next

Dataset Title	Agency	Categories	Catalog	Country
Elektronischer Wasserstraßen-Informationsservice - [Karte Nordwest]	Niedersächsischer Landesbetrieb für Wasserwirtschaft K sten- und Naturschutz	Wasser	Portalu.de (Germany)	Germany
Hochwasserrückhaltebecken online	Niedersächsischer Landesbetrieb für Wasserwirtschaft K sten- und Naturschutz	Wasser	Portalu.de (Germany)	Germany
Bergbau - Kontrollergebnisse	Landesamt für Geologie und Bergwesen Sachsen-Anhalt LAGB	Weitere	Portalu.de (Germany)	Germany
Lufthygienische Überwachung Schleswig Holstein	Lufthygienische Überwachung Schleswig-Holstein	Luft	Portalu.de (Germany)	Germany
Überwachung der Umweltradioaktivität (IMIS)	Bundesamt für Strahlenschutz	Strahlung	Portalu.de (Germany)	Germany
Hochwasserinfos für Brandenburg	Ministerium für Umwelt Gesundheit und Verbraucherschutz des Landes Brandenburg	Wasser	Portalu.de (Germany)	Germany
Aktuelle Pegel der Elbe in Brandenburg	Ministerium für Umwelt Gesundheit und Verbraucherschutz des Landes Brandenburg	Wasser	Portalu.de (Germany)	Germany
	chemikalien, gesundheit			

Catalogs: Deselect All

- (395459) Data.gov --- Geodata Catalog
- (7702) Data.gov.uk (United Kingdom)
- (5978) Data.gov.sg (Singapore)
- (3958) OpenDataCordoba.com(Cordoba,Spain)
- (3707) Data.gov (United States) --- Raw Data
- (2489) Data.gov.bc.ca (British Columbia, Canada)
- (1616) Data.govt.nz (New Zealand)
- (1165) Data.medicare.gov (USA)

Countries: Deselect All

- (406280) United States
- (8940) United Kingdom
- (5978) Singapore
- (5907) Spain
- (4777) Canada
- (1687) New Zealand
- (1293) Italy
- (776) Australia

Agencies: Deselect All

- (2223) Ayuntamiento de Córdoba
- (1666) Environmental Protection Agency
- (1468) Health
- (1347) Economic Development Board
- (1123) Department of Health
- (1117) U.S. Geological Survey Earth Resource
- (883) Land Information New Zealand
- (781) Department for Communities and Local Government
- (758) Regione Autonoma della Sardegna

Categories: Deselect All

- (140) Datos Demográficos
- (139) Empleo
- (139) Infraestructuras
- (136) Labour
- (134) Finance
- (129) Sciences
- (126) Transportation and Storage
- (126) State sector performance

- Our metadata model is described [here](#).
- An overview of our procedure for getting and publishing metadata is described [here](#).
- **NEW!** XML/RDF "dump" of IOGDS dataset catalog downloadable [here](#) (6 GB)
- **NEW!** Turtle-formatted "dump" of IOGDS dataset catalog downloadable [here](#) (1 GB)

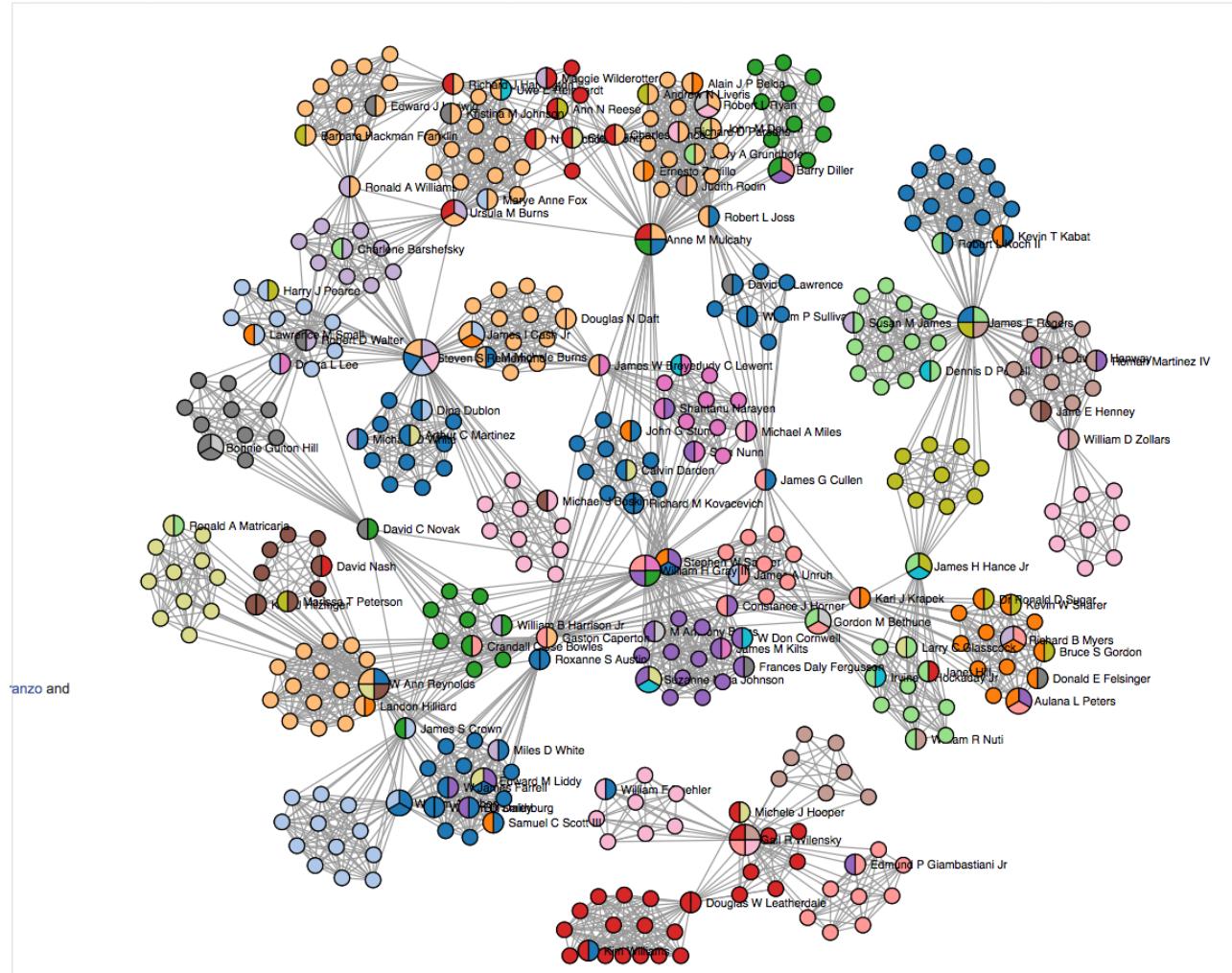
http://logd.tw.rpi.edu/demo/international_dataset_catalog_search

- Trust
 - ◆ Government data is controversial, and potentially biased
 - How do we confirm or dispute?
- Combination
 - ◆ When we combine data we need to keep the provenance of information (see trust)
 - How can we show and use?
- Scaling
 - ◆ Our project has already converted 9.9B triples from only >2,000 of the 440,000 government databases we can identify (116 catalogs, 38 countries, 16 languages)
- Versioning and updating
- Archiving
- Visualization
- ...

Exploring new visualizations

Board Members Network

Colored clusters represent boards of various US companies. Mouse-over nodes to see board member information. Click nodes to go to member's LittleSis page. Multi-colored nodes denote board members that are on multiple boards.



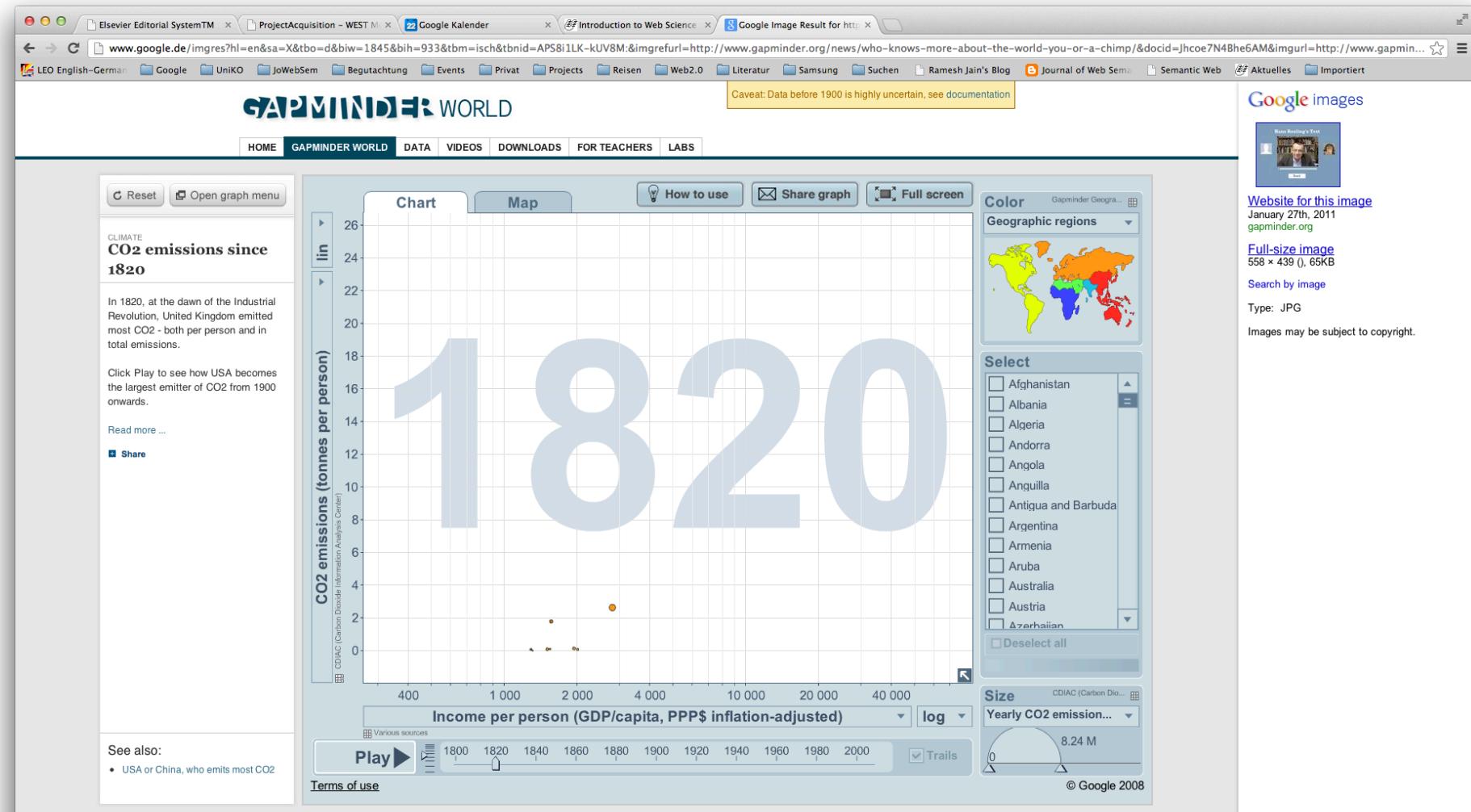
- Open Govt data is a critical resource
 - ◆ Government data released as RDF (UK)
 - ◆ Government data converted to RDF (US)
 - ◆ Government data that can be found in many forms and used or converted (WWW)
- Government transparency comes through in the “mashing up” of data from many datasets
 - ◆ Key to linked data
- An amazing opportunity for technologists (public and private) to play in an important area of the public good
 - ◆ Innovation needed!

EPILOG

Screenshot of a web browser showing the TED talk page for Hans Rosling's presentation "Stats that reshape your worldview".

The page includes the following elements:

- Header:** TED logo, navigation links (Talks, Speakers, Playlists, Translations, TED Conferences, TEDx Events, TED Prize, TED Ed, TED Conversations, TED Community, TED Fellows, TED Blog, TED Initiatives), and a search bar.
- Title:** TALKS, Hans Rosling: Stats that reshape your worldview.
- Meta Information:** FILMED FEB 2006 • POSTED JUN 2006 • TED2006
- Image:** A large thumbnail image of the TED stage with the TED logo and the text "TED Ideas worth spreading".
- Statistics:** 4,672,651 Views, Gefällt mir.
- Description:** You've never seen data presented like this. With the drama and urgency of a sportscaster, statistics guru Hans Rosling debunks myths about the so-called "developing world." In Hans Rosling's hands, data sings. Global trends in health and economics come to vivid life. And the big picture of global development—with some surprisingly good news—snaps into sharp focus. [Full bio »](#)
- Related Playlists:** Bill Gates: My 13 favorite talks (Curated by Bill Gates), Tim O'Reilly: 9 talks that stretched my mind (Curated by Tim O'Reilly), Steve Case: 10 unforgettable talks (Curated by Steve Case).
- TED Conversations:** Got an idea, question, or debate inspired by this talk? [Start a TED Conversation.](#)
- Comment Section:** COMMENT ON THIS TALK (399 total comments). Sign in to add comments or Join (It's free and fast!). Sort By: Newest first. A comment from Taivo Pungas is shown.
- What to Watch Next:** Hans Rosling: New insights on poverty (18:57 Posted: Jun 2007 Views 2,001,463 | Comments 212) and Paul Collier on the "bottom billion".



Now supported by 7% of web pages crawled by google

About 2% by those crawled by Yandex

Richer ontology just included (Good Relations)

No namespaces

„we will support what is popular“ Peter Mika, Yahoo

Winner of the Semantic Web Challenge 2012



The screenshot shows a web browser window with multiple tabs open, including "Elsevier Editorial SystemTM", "ProjectAcquisition - WEST M", "Google Kalender", "Introduction to Web Science", "Google Image Result for http://", "Raphael Troncy - semanticweb", "EventMedia", and "eventmedia.eurecom.fr/#/event/da356e86-2f1f-4612-996f-dec226fd06a3". The main content area is a dark-themed page for an event titled "Poets of the Fall at BLUE SHELL".
The event details are as follows:
Category: Musical Concert
Tags: Fall, Poets, concert, konzert, music, musik, rock, the, tickets, viagogo.com
When: Wed, 21 Nov 2012 20:00:00
16 HOURS AGO
Where: BLUE SHELL
Luxemburger Str 32
Köln, Nordrhein-Westfalen
Germany
Links: useful links
<http://www.viagogo.nl/Concertkaartjes/Rock+Pop/Poe...>
A map of Cologne (Köln) shows the location of the event at the Blue Shell venue. A photo of the band members is also displayed.
Below the event details, there is a section titled "Agents" which contains information about the band "POETS OF THE FALL":
"Without silence there can be no music." Poets of the Fall Finnish rockers Poets of the Fall - singer Marko, guitarist Olli and keyboardist...
At the bottom of the page, a copyright notice reads: © 2012, Eurecom, Sophia-Antipolis.