

0110
011
01
0



International Open Government Data Conference

The Semantic Web—What is it? Is it Ready for Prime-Time? And How Can I Use it?

Sir Tim Berners-Lee, *Director, World Wide Web Consortium;
Professor, Massachusetts Institute of Technology*



James A. Hendler, Ph.D., *Tetherless World Senior Constellation
Professor, Department of Computer Science and Cognitive
Science Department, Rensselaer Polytechnic Institute (RPI)*



George Thomas, *Enterprise Architect, U.S. Department of
Health and Human Services*

Marion A. Royal, *Agency Expert and Program Director,
Data.gov, Office of Citizens Services and Innovative
Technologies, U.S. General Services Administration
(Moderator)*



Government Open Linked Data

IOGDC 2010

[George Thomas, HHS](#)

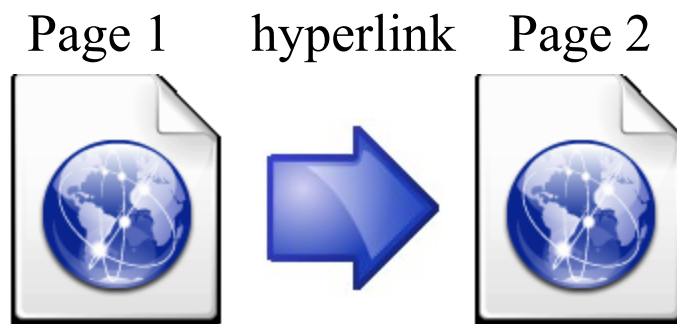
Data.gov PMO, Semantic Web and Linked Data lead

Agenda

- Intro, ‘What is Linked Data?’
 - From Webs of Docs, to Webs of Data
 - Examples from DBPedia and data.gov.uk
- Action Items
 - Vocabularies and URI Schemes
- RPI Tetherless World collaboration
 - Semantic mashup patterns, demos, and more
- Future Big Ideas

The Document Web

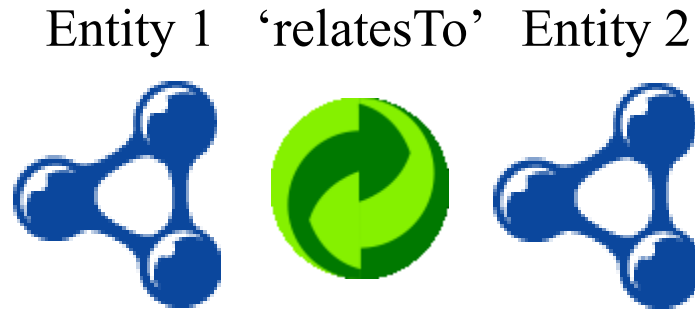
- Using HTML, there is only 1 *type* of link
 - An important 20th century invention for info workers!



- My page contains markup that links to Your page
 - `http://myserver.tld/mypage.html`
 - `link text`

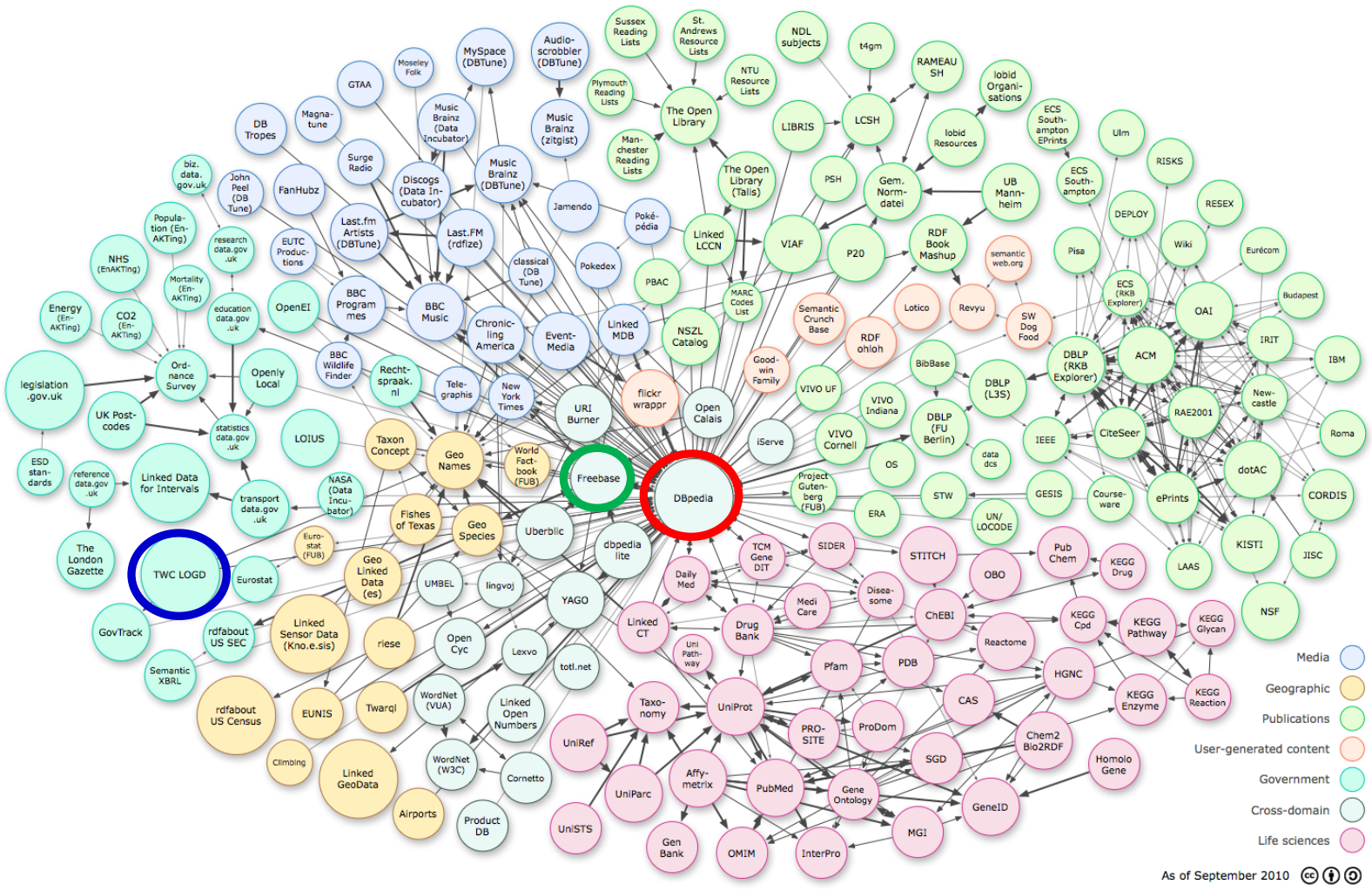
The Data Web


- Using RDF, we create *custom* ‘tagged’ links
 - An innovation where things relate to other things



- My data representation contains *triple* markup
 - `http://myserver.tld/Subject`
`<somePredicate> <http://yourserver.tld/Object>`.

Linked Data Cloud

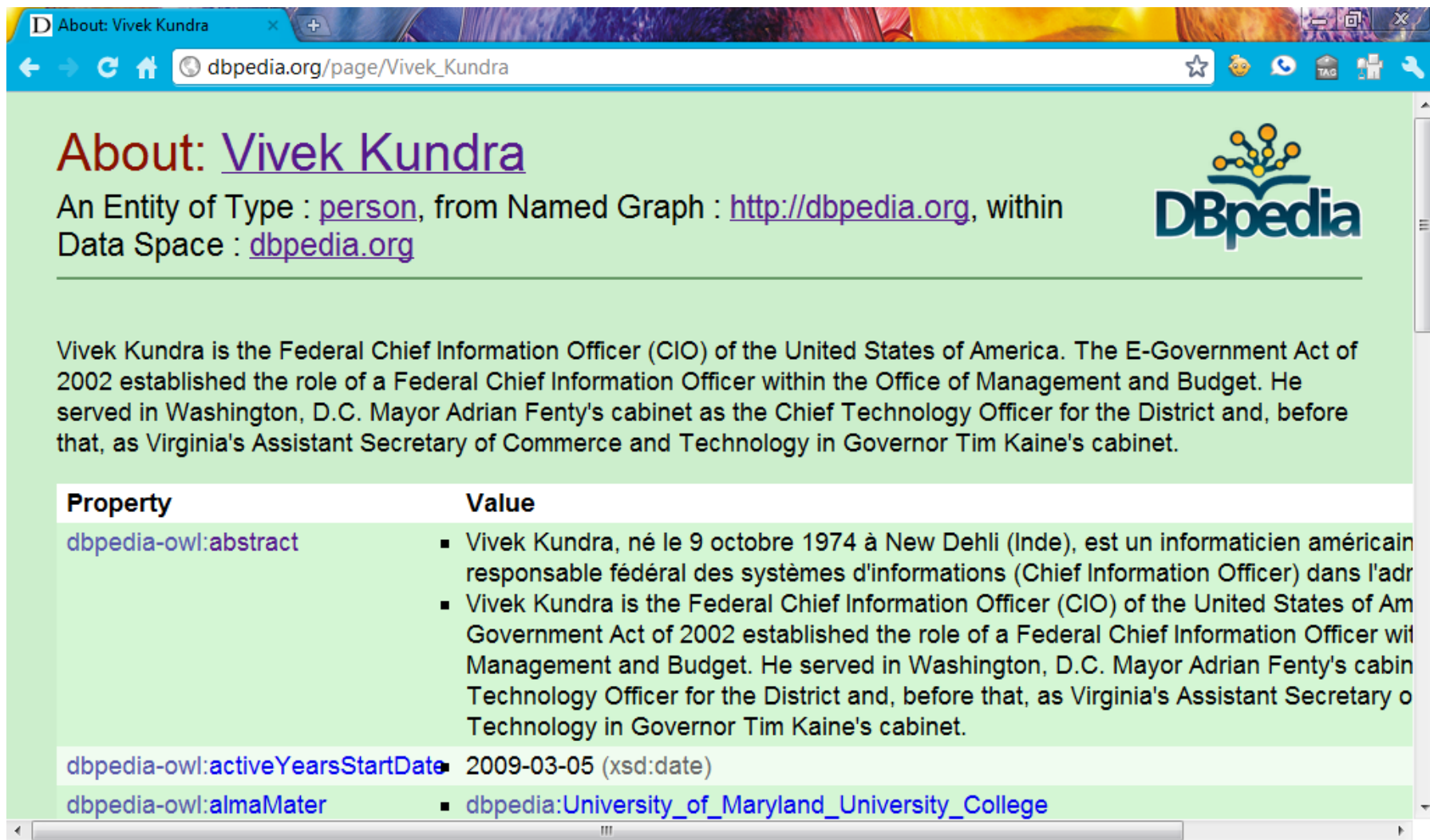


As of September 2010 

Content Negotiation

- Non-Information Resource (NIR)
 - http://dbpedia.org/resource/Vivek_Kundra
 - An HTTP dereferenceable (GET'able) URI for Mr. Kundra
 - HTTP Response Code, 303 'See Other'
- Redirects to an Information Resource (IR)
 - http://dbpedia.org/page/Vivek_Kundra
 - An HTML page 'about' Mr. Kundra
- Machine friendly representations (serializations)
 - http://dbpedia.org/data/Vivek_Kundra.rdf
 - Client specifies 'Accept' preference (mime-type) in request
 - Content Location points to representation URL

Basic Data Web Page



The screenshot shows a web browser window with the address bar containing `dbpedia.org/page/Vivek_Kundra`. The page title is "About: Vivek Kundra". The main content area has a light green background and features the DBpedia logo in the top right corner. The text on the page reads: "An Entity of Type : [person](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](#)". Below this is a paragraph of biographical information: "Vivek Kundra is the Federal Chief Information Officer (CIO) of the United States of America. The E-Government Act of 2002 established the role of a Federal Chief Information Officer within the Office of Management and Budget. He served in Washington, D.C. Mayor Adrian Fenty's cabinet as the Chief Technology Officer for the District and, before that, as Virginia's Assistant Secretary of Commerce and Technology in Governor Tim Kaine's cabinet." At the bottom, there is a table with two columns: "Property" and "Value".

Property	Value
<code>dbpedia-owl:abstract</code>	<ul style="list-style-type: none"> Vivek Kundra, né le 9 octobre 1974 à New Dehli (Inde), est un informaticien américain responsable fédéral des systèmes d'informations (Chief Information Officer) dans l'adr Vivek Kundra is the Federal Chief Information Officer (CIO) of the United States of Am Government Act of 2002 established the role of a Federal Chief Information Officer wil Management and Budget. He served in Washington, D.C. Mayor Adrian Fenty's cabin Technology Officer for the District and, before that, as Virginia's Assistant Secretary o Technology in Governor Tim Kaine's cabinet.
<code>dbpedia-owl:activeYearsStartDate</code>	2009-03-05 (xsd:date)
<code>dbpedia-owl:almaMater</code>	<code>dbpedia:University_of_Maryland_University_College</code>

A different representation...



```
{
  http://en.wikipedia.org/wiki/Vivek\_Kundra: {
    http://xmlns.com/foaf/0.1/primaryTopic: [
      {
        type: "uri",
        value: "http://dbpedia.org/resource/Vivek\_Kundra"
      }
    ]
  },
  http://dbpedia.org/resource/Vivek\_Kundra: {
    http://www.w3.org/1999/02/22-rdf-syntax-ns#type: [
      {
        type: "uri",
        value: "http://www.w3.org/2002/07/owl#Thing"
      },
      {
        type: "uri",
        value: "http://dbpedia.org/ontology/Person"
      },
      {
        type: "uri",
        value: "http://dbpedia.org/ontology/OfficeHolder"
      }
    ]
  }
}
```

Linked Data Principles

1. Use URI's as names for things
2. Use HTTP URI's so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards
4. Include links to other URI's so that they can discover more things

DBPedia - UK URI Set

- UK URI Set conventions similar to DBPedia mechanics, with different naming guidelines
 - (notional, just using VK as an example ;)

dbpedia.org	data.gov.uk
dbpedia.org/resource/Vivek_Kundra	data.gov.uk/id/Vivek_Kundra
dbpedia.org/page/Vivek_Kundra	data.gov.uk/doc/Vivek_Kundra
dbpedia.org/ontology/Person	data.gov.uk/def/{scheme}/Person
	{sector}.data.gov.uk/

Data Web Browser

OpenLink Data Explorer

linkeddata.uriburner.com/ode/?uri=http://dbpedia.org/resource/Vivek_Kundra

OPENLINK SOFTWARE Data Explorer

Data Source URI Go Find

What Where When Who Images Grid view Tag Cloud SVG Graph Navigator Custom

This view shows all RDF data grouped by subject resource.

Cache Total 42 triples All: [Check](#) [Uncheck](#) [Invert Sel](#) [Purge](#) [Refresh](#) [Permalink](#)

Vivek Kundra - 42 triples - [Remove](#) - [Refresh](#) - [Permalink](#)

There are 40 triples available.
[visible columns](#)

#	Subject	Predicate	Object
1	dbpedia:Vivek Kundra	dbpprop:reference	it.usaspending.gov
2	dbpedia:Vivek Kundra	dbpprop:reference	www.appsfordemocracy.org
3	dbpedia:Vivek Kundra	dbpprop:reference	start_page.do
4	dbpedia:Vivek Kundra	dbpprop:reference	www.USASpending.gov
5	dbpedia:Vivek Kundra	dbpprop:reference	federal cio viv.html
6	dbpedia:Vivek Kundra	dbpprop:reference	624870
7	dbpedia:Vivek Kundra	dbpprop:reference	data.octo.dc.gov
8	dbpedia:Vivek Kundra	dbpprop:reference	http://www.ctovision.com/2008/11/vivek...-putting-it-in-the-public-domain.html
9	dbpedia:Vivek Kundra	skos:subject	Category:Washington, D.C. Democrats
10	dbpedia:Vivek Kundra	skos:subject	Category:University of Maryland, College Park alumni
11	dbpedia:Vivek Kundra	skos:subject	category:1974 births
12	dbpedia:Vivek Kundra	skos:subject	category:Living people
13	dbpedia:Vivek Kundra	skos:subject	category:Obama Administration personnel
14	dbpedia:Vivek Kundra	skos:subject	category:Chief information officers
15	dbpedia:Vivek Kundra	skos:subject	category:American people of Indian descent
16	dbpedia:Vivek Kundra	skos:subject	category:Technology evangelists

- Categories
- Filters
- Previous Queries
- Bookmarks
- Data Retrieval Options
- Data Explorer Options

5 Star GOLD Rating

★ Make your stuff available on the Web
(whatever format)

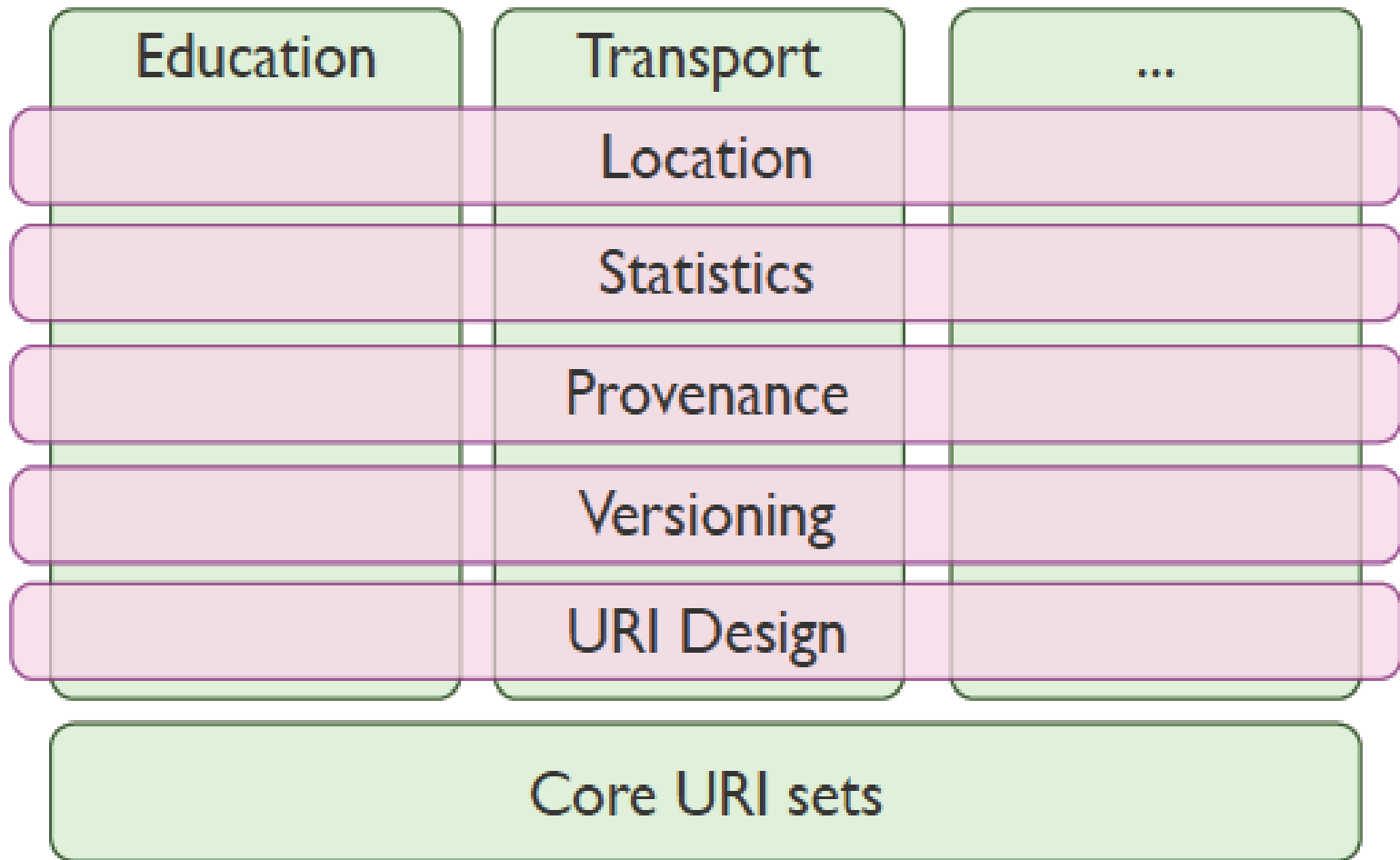
★★ Make it available as structured data (e.g. Excel
instead of image scan of a table)

★★★ Non-proprietary format (e.g. CSV instead of
Excel)

★★★★ Use URIs to identify things, so that people can
point at your stuff

★★★★★ Link your data to other people's data to
provide context

Vocabs and URI Schemes



Next Steps - Big Ideas

- Government Open Linked Data
 - Cross-cutting vocabularies become voluntary consensus stds
- Data Driven Journalism
 - ‘dataset of the day’, ‘#gplat as #dsub on the #ios’
- Data.gov provides (US) GOLD query service
 - Virtuoso ‘sponger’ ingest, ala linkeddata.uriburner.com
- Social Data Webs = Social Web + Data Web
 - Creating vocabularies, curating data
 - Socrata (HHS), Google Refine (data.gov.uk)
- Policy/Rule driven Data Entity Access Services
 - Rule Interchange Format
- [Linked-Data](#) API (open source)
 - Projecting SPARQL on the ‘Plain Web’

RPI Collaboration

- RPI has done so much with data.gov datasets!
 - RDF conversion from CSV
 - See [URI for LOD](#) design approach
 - Semantic Mashup Pattern and open source code
 - PHP (SPARQL -> XSLT -> JSON -> GoogleViz)
 - OpenLinkSW Virtuoso Triple (quad) Store
 - ‘SPARQL Endpoint’ query service ingest
 - Lots of cool mashups and apps!
 - Led first Data.gov ‘Mash-A-Thon’
- Thanks to Professor Hendler and the TW Team!

Thank You!