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

- 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

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

Page 1   hyperlink   Page 2
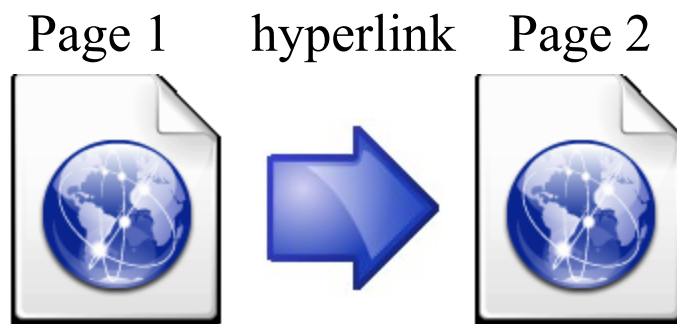
- My page contains markup that links to Your page
  - http://myserver.tld/mypage.html

    link text

- Using RDF, we create *custom* 'tagged' <u>links</u>
  - An innovation where things <u>relate</u> to other things

Entity 1   'relatesTo'   Entity 2
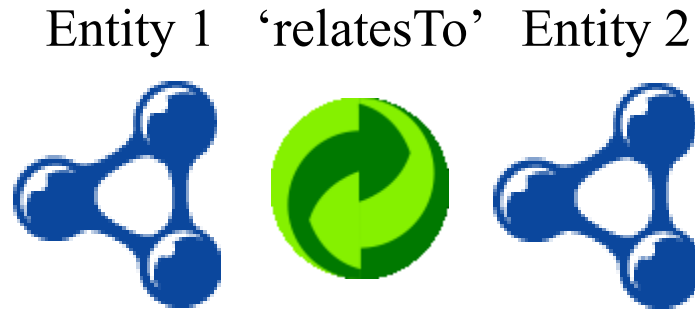
- My data representation contains ***triple*** markup
  - http://myserver.tld/Subject
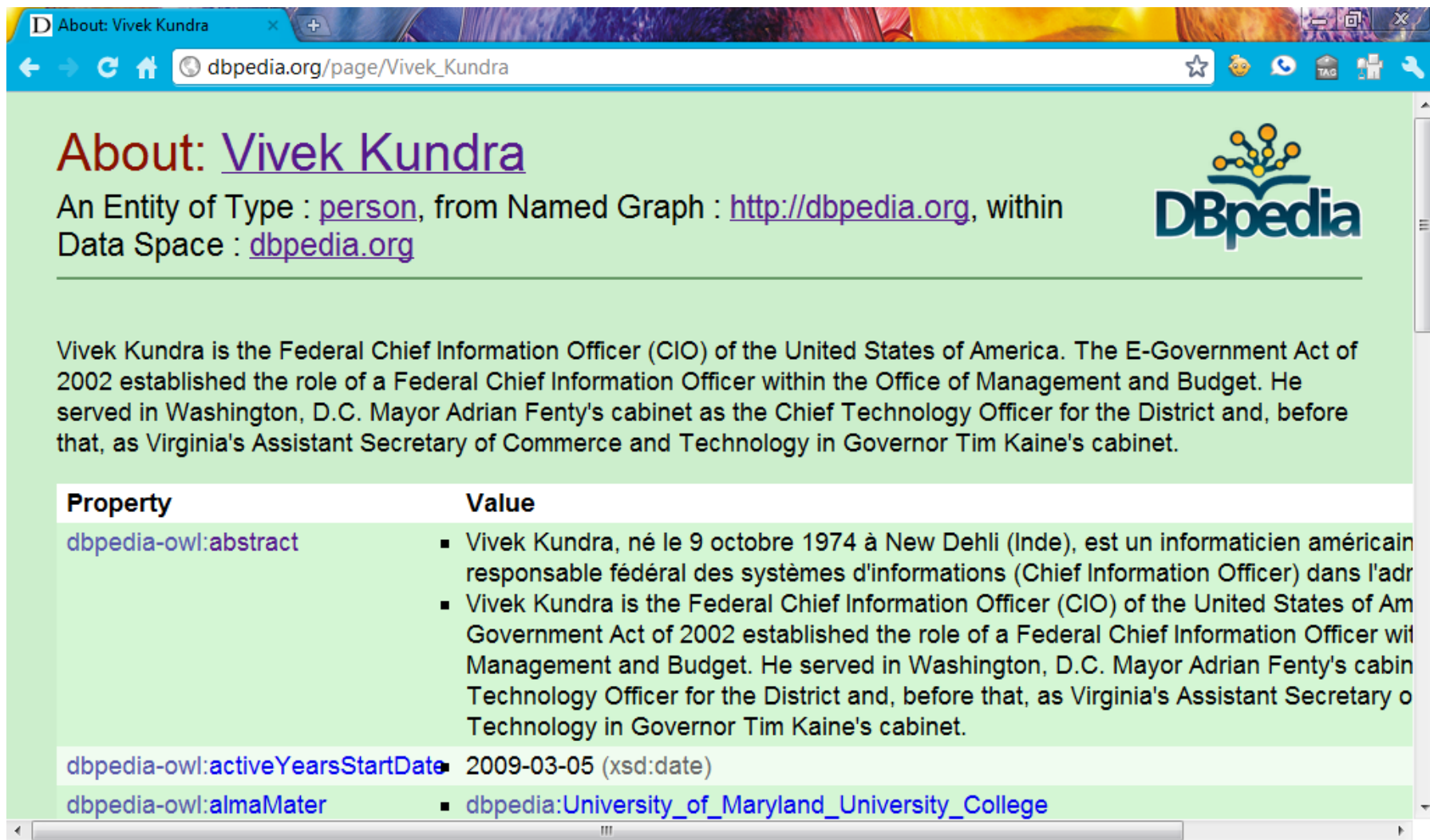   <somePredicate>   <http://yourserver.tld/Object>.

Linked Data Cloud
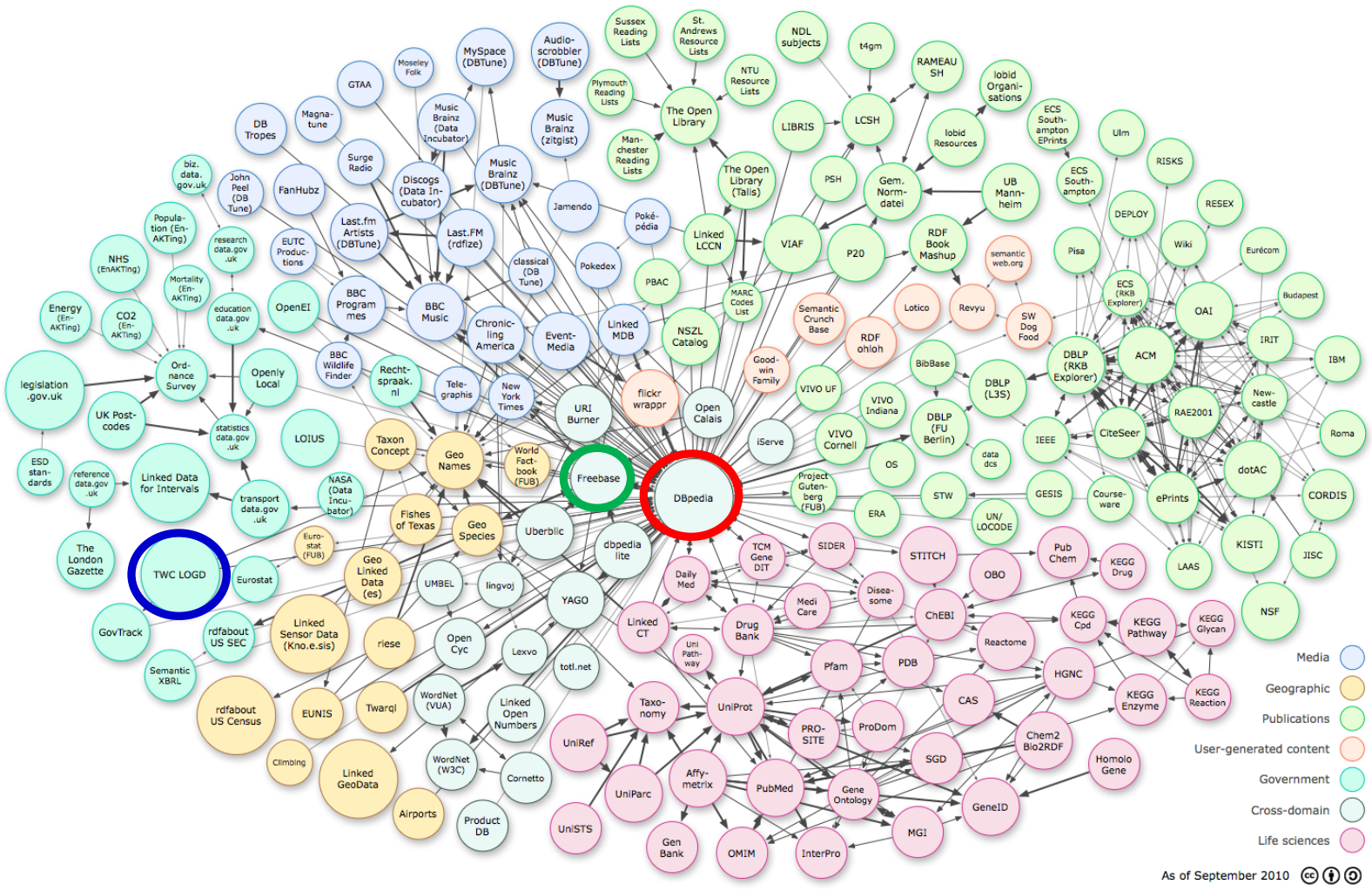
Content Negotiation

- 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

A different representation…

dbpedia.org/data/Vivek_...

dbpedia.org/data/Vivek_Kundra.json

```
{
    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

- ## 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

★ 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

Education	Transport	...
	Location	
	Statistics	
	Provenance	
	Versioning	
	URI Design	
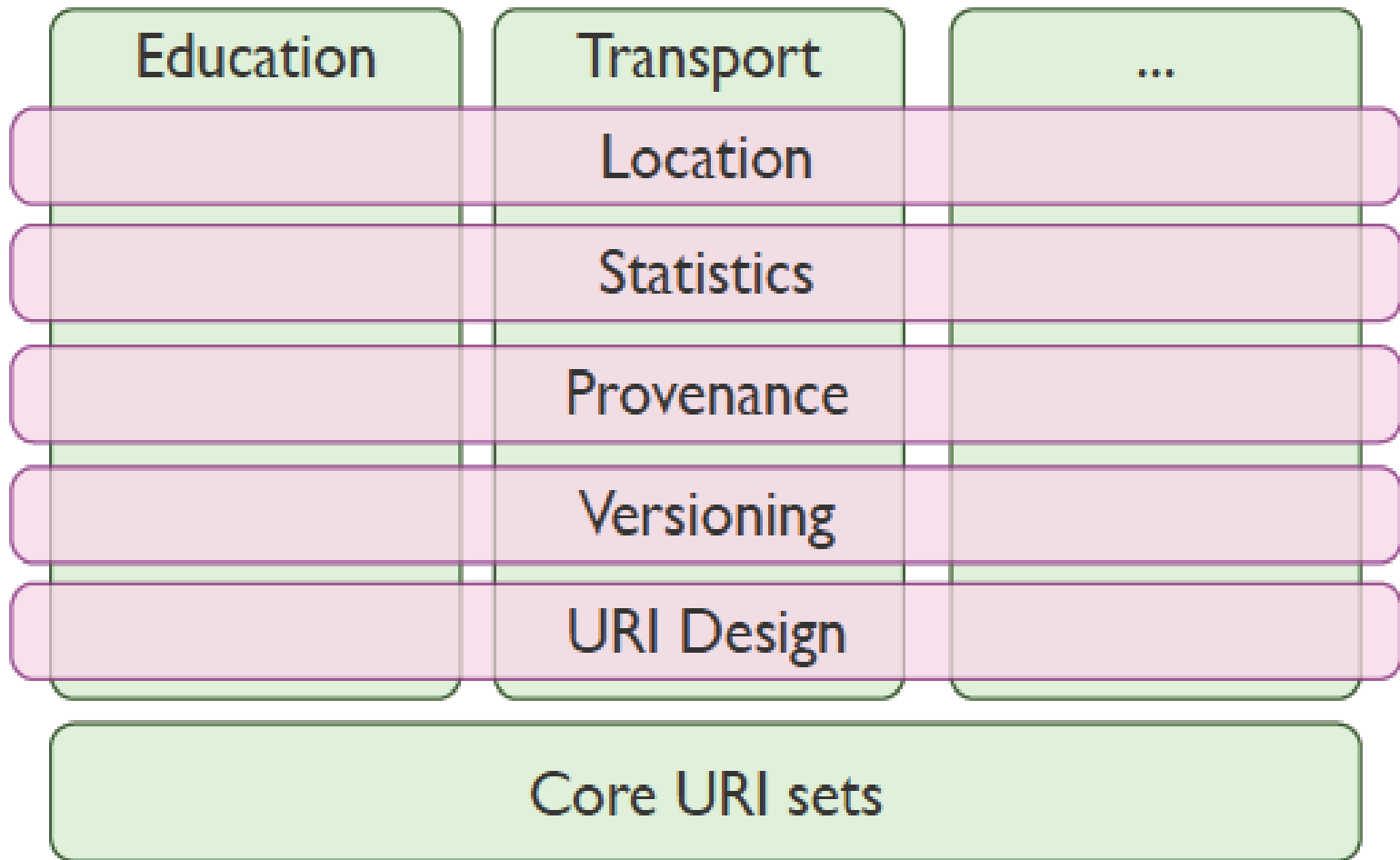
Core URI sets

- 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 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!