Semantic Search:
We’re Living in a Golden Age for Information

Bernadette Hyland, CEO & co-founder
co-chair W3C Government
Linked Data Working Group

NKOS-CENDI Semantic Search Workshop - Washington DC
6-December 2012
The Race to the Top for Openness and Innovation: Announcing Agency Open Government Plan Leading Practices

Big Data for Users/Leaders-Managers-Innovators-Investors and Asset Managers

October 26, 2012 By BobGourley

“The hottest tech trend of the year is Big Data.”

“... It has been a real revelation, finding a granular level of data that we never had before.”

Forbes
Big Data

Simple data
Complex data
Legacy data
“Big Data Is Important, but Open Data Is More Valuable”

As change agents, enterprise architects can help their organizations become richer through strategies such as open data.

David Newman, VP Research, Gartner
What am I publishing?

How can I structure it so it can be re-used?

How can I be more open?
Open Government Data

Steven VanRoekel @steenvDC
We need to get govt out of the presentation business and into the content and data business. wh.gov/digitalgov #digitalgov #tcdisrupt

Steven VanRoekel @steenvDC
It's time to make govt data social wh.gov/digitalgov #digitalgov #tcdisrupt

Steven VanRoekel @steenvDC
Govt data needs to be open, machine-readable and available in timely manner—by default. wh.gov/digitalgov #digitalgov #tcdisrupt

Steven VanRoekel @steenvDC
Everything should be an API wh.gov/digitalgov #digitalgov #tcdisrupt
Growing chorus ...

“We’re moving from managing documents to managing discrete pieces of open data and content which can be tagged, shared, secured, mashed up and presented in the way that is most useful for the consumer of that information.”

The Web of Data

- A global network of linked **statements**
- A place where anyone can say anything about anything
- A vast collection of **machine**-readable knowledge (and opinion)
- **Statements** are linked, and links are qualified
• Linked Data is about publishing and consuming data using international data standards

• Based on 20 year old idea

• A system of linked information systems
Governments

Goals: Governmental transparency and/or improved internal efficiencies (data warehouses)
Open data + open standards + open platforms

Highly scalable computing & hosting via the Cloud

International Data Exchange Standards

5 Star Data (Linked Data)

Open Source tools
Linked Data Improves Discoverability & Access

Linked Data - refers to a set of best practices for publishing and interlinking structured data for access by both humans and machines via the use of the RDF family of syntaxes (e.g., RDF/XML, N3, Turtle and N-Triples) and HTTP URIs.

Linked Data can be published by an person or organization behind the firewall or on the public Web. If Linked Data is published on the public Web, it is generally called Linked Open Data.

- W3C Linked Data Glossary: https://dvcs.w3.org/hg/gld/raw-file/default/glossary
Integrating ...

- Big data
  - Simple data
  - Complex data
  - Legacy data
THERE IS A PROCESS

Identify → Model → Name → Describe → Convert → Publish
THERE IS A PROCESS

Identify ➔ Model ➔ Name ➔ Describe ➔ Convert ➔ Publish

Maintain
We’ve Seen This Before
Callimachus is an open source, open standards platform for Web developers to create data-driven applications

http://callimachusproject.org
Contractor (3 Round Stones, Inc.)

Public

Application, Script or automated client

Linked Data management system located at a Tier 1 Cloud Provider (FISMA compliant)

RDF Database

Resource URIs

REST API

SPARQL endpoint

Web Browser

Application, Script or automated client

Linked Data management system located at a Tier 1 Cloud Provider (FISMA compliant)
Asthma

Asthma (from the Greek "panting") is the common chronic inflammatory disease of the airways characterized by variable and recurring symptoms, reversible airflow obstruction, and bronchospasm. Symptoms include wheezing, coughing, chest tightness, and shortness of breath. Asthma is clinically classified according to the frequency of symptoms, forced expiratory volume in 1 second, and peak expiratory flow rate. Asthma may also be classified as atopic (extrinsic) or non-atopic (intrinsic). It is thought to be caused by a combination of genetic and environmental factors. Treatment of acute symptoms is usually with an inhaled short-acting 
bronchodilator. Spacers are often used with long-acting bronchodilators.

Hydrocodone Acetaminophen

WARNINGS
Respiratory Depression: At high doses or in sensitive patients, hydrocodone may produce dose-related respiratory depression by acting directly on the brain stem respiratory center. Hydrocodone also affects the center that controls respiratory rhythm, and may produce irregular and periodic breathing. Head Injury and Increased Intracranial Pressure: The respiratory depressant effects of narcotics and their capacity to elevate cerebrospinal fluid pressure may be markedly exaggerated in the presence of head injury, other intracranial lesions or a preexisting increase in intracranial pressure. Furthermore, narcotics produce adverse reactions which may be markedly exaggerated in patients who are already impaired by other CNS depressant drugs.

Alerts

2 of 11

Category: Good

Enter a VA ZIP code above

Partly cloudy

59°

41%

49 pm2.5

2 of 11

Alert: False

Hydrocodone Acetaminophen

Enter a VA ZIP code above
Asthma

Asthma (from the Greek "panting") is the common chronic inflammatory disease of the airways characterized by variable and recurring symptoms, mostly coughing, chest tightness, and shortness of breath. Asthma is clinically classified according to the frequency of symptoms, forced expiratory volume in 1 second, and peak expiratory flow rate. Asthma may also be classified as atopic (extrinsic) or non-atopic (intrinsic). It is thought to be caused by a combination of genetic and environmental factors. Treatment of acute symptoms is usually with an inhaled short-acting beta-agonist, and long-term control is usually with an inhaled corticosteroid.
Data driven Web apps using Callimachus

US Legislation + enterprise data

DBpedia + enterprise datasets

Clinical Trials + enterprise linked data
Key points

• No ‘one size fits all’
  • Simple, complex and legacy data require different approaches
  • Search, discovery and access are coming together

• Open Standards
  • Linked Data improves search, discovery & access
  • HTML5 for great user experience, mobile

• Biggest win for government & the public is ease of combining data sets
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**

<table>
<thead>
<tr>
<th>Start Date</th>
<th>June 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>End date</td>
<td>31 May 2013</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>Proceedings are <strong>public</strong></td>
</tr>
<tr>
<td>Chairs</td>
<td>George Thomas, US HHS Bernadette Hyland</td>
</tr>
<tr>
<td>Staff Contact</td>
<td>Sandro Hawke (0.3 FTE)</td>
</tr>
<tr>
<td>Usual Meeting Schedule</td>
<td>Teleconferences: Weekly, plus optional task forces Face-to-face: 2-4 per year, possibly multisite with video</td>
</tr>
</tbody>
</table>
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.
• Additional information available on the Web, in books ...

http://linkeddatabook.com/editions/1.0/
http://3roundstones.com/linking-enterprise-data/
http://3roundstones.com/linking-government-data/
http://www.manning.com/dwood/
http://www.w3.org/2011/gld/wiki/
Linked_Data_Cookbook

• Commercially supported Open Source Web data platform

http://3roundstones.com/products

Bernadette Hyland
@BernHyland
bhyland@3roundstones.com

Based in Virginia
This work is Copyright © 2011-2012 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

<table>
<thead>
<tr>
<th>Author/Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>W3C Linked Data Cookbook</td>
<td><a href="http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook">http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook</a></td>
</tr>
</tbody>
</table>

All other photos and images © 2010-2012 3 Round Stones, Inc. and released under a CC-by-sa license