Delivering on Standards for Government Linked Data

W3C Working Group Report

Bernadette Hyland
co-chair, W3C Government Linked Data WG
CEO, 3 Round Stones, Inc.

@BernHyland
10-April-2013

European Data Forum 2013 Dublin

Context of today’s session is to briefly describe the activities & deliverables of the W3C Government Linked Data working group & what we’re doing.

Following the EU Data Forum, the GLD WG has its 3rd F2F meeting. This will be an interactive discussion about what each of us needs to more effectively work with government, as well as, review of feedback on the working group’s deliverables.
Mark Gerzon’s book *Global Citizen* introduces the concept of global intelligence and levels of citizenship, Citizen 1.0 thru 5.0. Citizen 4.0s are those with a worldview based on multiple cultures (multi-centric) and Citizen 5.0 infers a worldview based on a Whole Earth perspective (geocentric).
I’m speaking today on behalf of approximately 25 active international “bridge builders” -- We are members of W3C Government Linked Data WG. Note: I’m a volunteer working group chair, not a W3C staff member. This is a brief update on resources to help governments publish structured data on the Web.
What exactly is 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. Linked Data published on the public Web is generally called Linked Open Data.
What was the catalyst for the formation of this government data standards group? In short, to seek guidance on and share information about publishing structured content on the Web in a way in which it can be readily be found, accessed, shared and re-used with others.
Themes

**Myth:** Data standards are published and therefore government can jump on board publishing all kinds of useful data.

**Reality:** Formats for machine readable data are well-documented. However, modeling data with useful context, “pragmatic” provenance & visualizations is not trivial.

Plan for new approaches ... they take time.

Open data on the Web requires a change in mindset. On one hand IT & security people are being asked to change tools & approach, often from ‘need to know’ to ‘need to share’ mentality using Open Source tools with which they aren’t familiar. They can become ‘blockers’. In contrast, data stewards and regulators are often keen to widen the reach of their data but need technical guidance. Researchers in many cases have great recommendations however their role isn’t to implement & support high availability systems.
Themes

**Myth:** Our agency is providing everything on the Web via a portal backed by a central relational database.

**Reality:** On the Web (in HTML) does not mean machine readable. Also, most data is organized by programs or regulations that are not intuitive making access nearly impossible. Context is often non-existent making data nearly useless.

There is no question that today is more fashionable to promote open government data projects versus a proprietary €20M data warehouse project. That said, the fiscal cliff the US & EU face, combined with big data and deploying systems on the cloud have contract officers & managers flat out trying evaluating options that range from the status quo to disruptive.
Themes

**Myth:** We’re open!! We offer a data catalogue, publish APIs & our open government initiative is available (as a gorgeous PDF document)!

**Reality:** Yes, civically minded developers have delivered useful apps which is encouraging.

Yet governments spend billions of € collecting data that remains *inaccessible & unusable* to scientists, researchers & the public.

---

A ray of sunshine! Open government data advocates take great comfort that highly visible government executives, many of whom are here at this conference, support in words and with budgets, well-considered open government data projects. Thank you!! That said, the state-of-the-art is a decade beyond screen scraping. That is not the approach developers should have to adopt.
Themes

**Myth:** There is plenty of open government data available that Web developers are using to spur economic development.

**Reality:** Yes, there are >1M government datasets published on the Web, in nearly 200 catalogues, published by 43 countries, in 24 languages however, they are often hard to find & require considerable programming effort to use.

Human health & environmental scientists spend over 50% of their time massaging data and models to interoperate. If they had better data interoperability, the global R&D community would be significantly more efficient & effective. The stakes with controlling the cost & effectiveness of healthcare delivery and global warming are issues most governments face. That is why data interoperability really matters!

Source of statistics: RPI IOGDS Data Analytics Project Statistics as of 2013, see http://logd.tw.rpi.edu/iogds_data_analytics
Can someone please make information easier to consume??

The state of publishing data and making it accessible to humans and machines is evolving rapidly because the stakes are so high.

I'm happy to report that the state of the art for publishing structured data on the Web is ready for prime time -- ready for governments to seriously invest in for near and medium term benefits.
Yes!

The **mission** of the W3C [Government Linked Data (GLD) Working Group](https://www.w3.org/2013/GLD) is to provide standards and other information which help governments around the world publish their data as effective and usable [Linked Data](https://www.w3.org/2001/sw/wiki/LinkedData) using [Semantic Web](https://www.w3.org/2001/sw/wiki/SemanticWeb) technologies.

We are 22 months into the Government Linked Data Working group's two year charter. This is the beginning of a 5-10 year effort IMO. Should you start budgeting for (modeling) and publication of authoritative data as Linked Data? Absolutely YES!
How do we get a loose coupling of shared data over Web architectures? By using the structured data model for the Web: RDF.

There is a project to create freely available data on the Web in this way, which is known as the Linked Open Data project.

W3C sees Linked Data as the set of best practices and technologies to support worldwide data access, integration and creative re-use of authoritative data.
The **Web of Data** is helping to meet government missions

- A global database of linked statements
- A place where governments can publish authoritative information
- A vast collection of **both human & machine** readable information
- Agencies are beginning to publish Open Data as “5 star Linked Data”

Over the last decade, a Web-oriented approach to information sharing has impacted how scientists, researchers, regulators and the public interacts with government. Linked data lowers the barriers to re-use and interoperability among multiple, distributed and heterogeneous data sources.

Governments in the European Union, US, Australia, New Zealand, Brazil and China are defining digital strategies sets to accomplish: Access to high quality digital information & services; procure and manage devices, applications, and data in smart, secure and affordable ways; and unlock the power of government data to spur innovation.

Access to high-quality Linked Open Data via the Web means millions of researchers and developers will be able to shorten the time-consuming research process involving data cleansing and modeling.
Working Group Deliverables
See http://www.w3.org/2011/gld/wiki/Main_Page

- Best Practices - includes guidance on URI policy and persistence
- Vocabularies: Organization, Data Cube (for statistics), DCAT (data catalog)
- A Linked Data Cook Book
- Community Directory
- Working Group Notes
  - Glossary (Linked Data terms)
  - ADMS (Asset Description Metadata Schema)

Support is available to government stakeholders and their contractors from the W3C and working groups such as the Government Linked Data WG, and others: Providence, RDF, Linked Data Platforms.
Government assumes responsibility ...

• Governments collect important and often expensive information

• Scientists, researchers, policy makers & the public are depending on the publication of reliable open government data

• As the authority, you have the responsibility to ensure information doesn’t disappear ...

  Just say no to “Page not Found: 404”
Introduction

The Community Directory contains a listing of organizations creating and/or publishing data on the Web. The Directory is open and does not require W3C affiliation. The directory has recently been restructured to be more "Web-like" and serves as an aggregator of Linked Data published on the Web. Add your organization's information to the Directory by following these easy steps.

Flow Diagram

1. Use the Directory form to create an entry
2. Download the generated file and publish it on the Web
3. Submit the URL to the Directory
4. The Directory retrieves and publishes your organization's information

Steps to Add Organization

Step #1 - Create an entry which generates RDF for your organization.
Step #2 - Download the file and publish it on the Web.
Step #3 - Submit the URL of the file containing your organization's listing.
Step #4 - The directory retrieves and publishes your organization's entry.

See the Directory Listing and Statistics to view existing entries.

Disclaimer: The W3C has not performed any evaluation and does express any judgment about the information provided. If any material in the directory is contested, the W3C may decide to cite and quote third-party refutations.
STATISTICS

Organizations by Country

Projects by Country

Projects by Type of Organization

Projects by Industry or Domain Category

- Commercial
- Research / Academic
- Federal / Central Govt
- Non-Profit Organizations
- Local Government
- State Government

- Government Agency or Department
  - Visualization
  - Internet S
  - Consulting
  - Healthcar
  - 48.2%
http://dir.w3.org

- If your government is publishing Linked Open Data, please add it to the *new and improved* W3C Directory!

- 100% open Web standards-compliant

- Built on an open source Linked Data Platform

- As easy as 1, 2, 3!!
Navigating Open Government Data need not require experimentation

Open Web standards allow us to use simple and complex data to make a real difference in people’s lives.

**Government Linked Data can help you navigate**

[http://www.w3.org/2011/gld/wiki/Main_Page](http://www.w3.org/2011/gld/wiki/Main_Page)

*Photo: Amelia Earhart, American aviation pioneer & author*

The W3C Government Linked Data WG’s deliverables address the key issues that motivate government to invest in improving the worldwide INFORMATION infrastructure, the Web of Data. It is early days, but the weather is fine and we encourage you to participate in whatever way you can with use cases, technical, marketing and communications expertise.
http://slideshare.com/3roundstones

Twitter: @BernHyland
Email: bhyland@3roundstones.com

Thank you for participating!!
Info for developers

http://www.manning.com/dwood/

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

http://3roundstones.com/linking-enterprise-data/
This presentation prepared by Bernadette Hyland, 3 Round Stones, Inc. is licensed under a Creative Commons BY-SA license, allowing you to share and remix its contents as long as you give us attribution and share alike.