Getting to 'the 5 stars of Linked Open Data' for Nanoinformatics

Mills Davis, Project10x, and Brand Niemann, Semanticommunity.net

November 4, 2010
Need: A database to keep track of this and everything else in nanotechnology/informatics!

Goal: Provide Semantic Data in each of our Semantic Community Wikis using Spotfire and Concept-map Ontology Environment softwares.

http://www.whitehouse.gov/search/site/nanotechnology
Manual extraction of web pages to a spreadsheet – like to automate, but sometimes it is too difficult to find a pattern or time-consuming to write the code.

http://semanticommunity.wik.is/@api/deki/files/1750/=Nanoinformatics.xls
Semantic Community.net


http://semanticommunity.net/
Abstract

- Tim Berners-Lee has suggested recently implementing Open Linked Data as part of a continuum of web publishing activities associated with gold stars, like the ones you got in school, as follows: 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; and link your data to other people’s data to provide context.

- This presentation will illustrate and demonstrate each of these 5 stars with the Nanoinformatics 2007 and 2010 Workshop content using Wiki, business intelligence analytics and visualization, and concept-map ontology environment softwares.
Background

• June 13, 2007:
  – Nanoinformatics: Locate, Collaborate, and Integrate, Brand Niemann, EPA, and Mills Davis, Project10x.
  • Slides.
  • Links: SICoP (Wiki), Nanoinformatics Deki Wiki Pilot (see next page), Nanoinformatics Semantic Wiki Pilot (decommissioned), and Nanoiformatics 2007 Knowledgebase Pilot (structured, but not well-defined URLs).
Nanoinformatics 2007 in a Wiki

http://semanticommunity.wik.is/Nanoinformatics
Nanoinformatics 2007 in a Wiki

http://semanticommunity.wik.is/Nanoinformatics/Participants/Mark_Tuominen
The Five Stars of Linked Open Data

The 5 stars of open linked data

While perusing the minutes of today’s w3c day telecon I noticed mention of Tim Berners-Lee’s Bag of Chips talk at the gov2.0 expo last week in Washington, DC. I actually enjoyed the talk not so much for the bag-of-chips example (which is good), but for the examination of Linked Data as part of a continuum of web publishing activities associated with gold stars, like the ones you got in school. Here they are:

* 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

I think it’s helpful to think of Linked Data in this context, and not to minimize (or trivialize) the effort and the importance of getting the first 3 stars.

It was interesting that he didn’t mention RDF once (unless I missed it) and talked instead about Linked Data Format. Correction he did mention it, thanks Andy. The indiscernability and ambiguity appeals to me.

http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/
The Five Stars of Linked Open Data

• 1. Make your stuff available on the web (whatever format);
• 2. Make it available as structured data (e.g. excel instead of image scan of a table);
• 3. Non-proprietary format (e.g. csv instead of excel);
• 4. Use URLs to identify things, so that people can point at your stuff; and
• 5. Link your data to other people’s data to provide context.
How to Publish Linked Data on the Web

Authors:
Chris Bizer (Web-based Systems Group, Freie Universit\aet Berlin, Germany)
Richard Cyganiak (Web-based Systems Group, Freie Universit\aet Berlin, Germany)
Tom Heath (Knowledge Media Institute, The Open University, Milton Keynes, UK)

This version:
http://sites.wiwiss.fu-berlin.de/pub/bizer/pub/linkedDataTutorial/20070727/

Latest version:
http://sites.wiwiss.fu-berlin.de/pub/bizer/pub/linkedDataTutorial/

Abstract
This document provides a tutorial on how to publish Linked Data on the Web. After a general overview of the concept of Linked Data, we describe several practical recipes for publishing information as Linked Data on the Web.

Table of Contents
1 Introduction: Linked Data on the Web
2 Basic Principles
  1 Web Architecture
  2 The RDF Data Model
3 Choosing URIs
4 Which vocabularies should I use to represent information?
5 What should I return as RDF description for a URI?

http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/
Semantic Wikis

Annotate in a Semantic Wiki on the way to say a DBpedia.

http://networkcentricity.wik.is/Semantic_Wikis
1. Make your stuff available on the web

http://www.nanotechinformatics.org/
2. Make it available as structured data

http://semanticommunity.wik.is/@api/deki/files/1750/=Nanoinformatics.xls
3. Non-proprietary format

- A. Spotfire Imports Excel, etc. and Exports to CSV:
  - Silver Free – one year free
  - Professional – 30 day free
- B. Concept-mapping Ontology Environment (COE) Imports Concept Maps, OWL, etc. and Exports RDF/OWL:
  - Free Desktop Software
  - Free Share Folder on the Florida Institute for Human and Machine Cognition Server for World-wide sharing
3A. TIBCO Spotfire Silver

One year free trial with only 3 files of less than 10 MB each

- Publish personal and professional dashboards & analytic applications
- No time-consuming setup or infrastructure required
- Easy analytics exchange and collaboration

The TIBCO Silver™ Spotfire® program is available for an initial one-year trial and provides an authoring client, expansive Web-based sharing, and hosting for your favorite personal or business Spotfire application.

TIBCO Silver Spotfire makes business intelligence and analytics more accessible through its social networking-like experience, allowing you to interact with data and share results in one click over the Web. Teams and workgroups can exchange ideas and insights and online publishers can embed live BI dashboards and reports into their websites or blogs.

Register for your one-year trial now:

1. Import your favorite data
2. Build a BI dashboard or report
3. Instantly publish it to our cloud servers

Try TIBCO Silver Spotfire today, with no cost or obligation!

See: http://spotfire.tibco.com/silverspotfire/faq.aspx

http://goto.spotfire.com/g/?KXLMSEWRM9
3A. TIBCO Spotfire Professional

30-day free trial with unlimited number of files and file sizes

3B. IHMC Cmap Tools

http://cmap.ihmc.us/
3B. Florida Institute for Human and Machine Cognition COE

http://www.ihmc.us/groups/coe/
4. Use URLs to identify things, so that people can point at your stuff

- Use MindTouch Wiki to Chunk Information by Page or Sub-Page (slide 12)
- Put in Excel Spreadsheet (slide 14)
- Make Web Links Active in Spotfire (coming later)
- Part of Ontology in COE (coming soon)
5. Link your data to other people’s data to provide context

• Two or more Excel tables linked in Spotfire and COE are what you need to get started.
  – Rebecca Reznik-Zellen helped us get started (see next slide).

• Spotfire supports Linked Open Data, but not using RDF at present.
  – This interface is being worked on by Spotfire and Kingsley Idehen, Founder and CEO of Open Link Software.
caBIG Nanotechnology Working Group Nanotechnology resources

http://sites.google.com/site/cabignanowg/nanotechnology-resources
Nano Data in Excel

NanoParticle Ontology for Cancer Nanotechnology Research

http://www.nano-ontology.org/
Build Nonoinformatics with Semantic Cloud Computing Tools

http://semanticommunity.wik.is/Nanoinformatics_2010
Build Nonoinformatics with Semantic Cloud Computing Tools
Some Next Steps

• Continue with some more manual extraction of key Nanoinformatics web pages to a spreadsheet.
• Use (or write) simple codes to automatically extract key Nanoinformatics databases to CSV.
• Inventory key Nanoinformatics databases in CSV (Excel) and import them into Spotfire for Linked Data Analytics.
• Continue to inventory Nanoinformatics ontologies and import them into COE and develop new Nanoinformatics ontologies to support improve search and reasoning across diverse Nanoinformatics content.
• And of course implement your suggestions and report back at the next Informatics Meeting!
Extra Slides of Our Work

• Nanomaterial-Biological Interactions Knowledgebase
• InterNano Taxonomy: Top Level
• InterNano Taxonomy: Bottom Level
• IntreNano Taxonomy: Spreadsheet
Nanomaterial-Biological Interactions Knowledgebase

http://nbi.oregonstate.edu/knowledgebase/analysis
InterNano Taxonomy: Top Level

http://eprints.internano.org/view/internano_taxonomy/
InterNano Taxonomy: Bottom Level

http://eprints.internano.org/view/internano_taxonomy/internanotaxonomy5827/
IntreNano Taxonomy: Spreadsheet

http://semanticommunity.wik.is/@api/deki/files/1750/=Nanoinformatics.xls

Annotate with your comments.