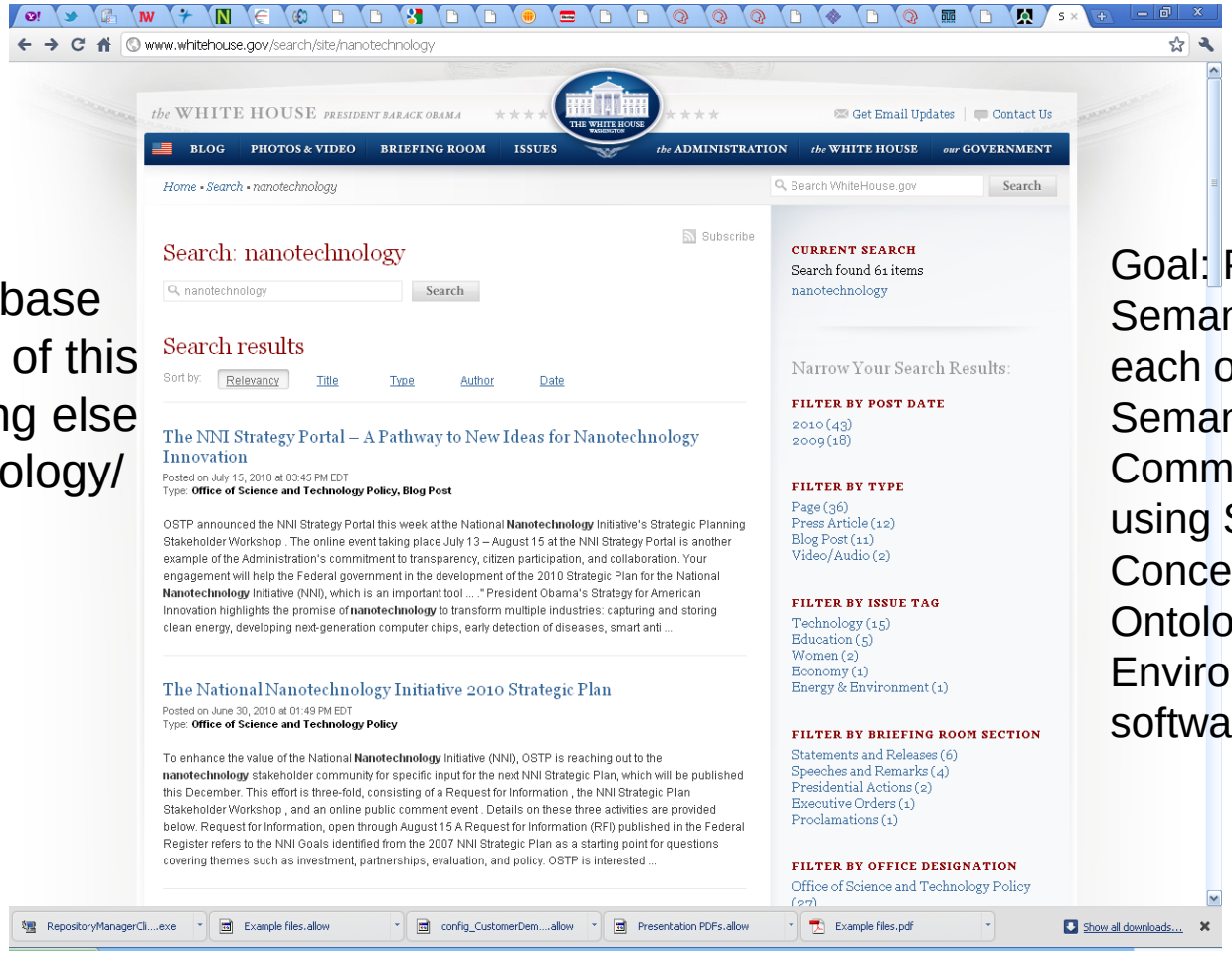
Getting to 'the 5 stars of Linked Open Data' for Nanoinformatics

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

November 4, 2010

White House Blog: Nanotechnology

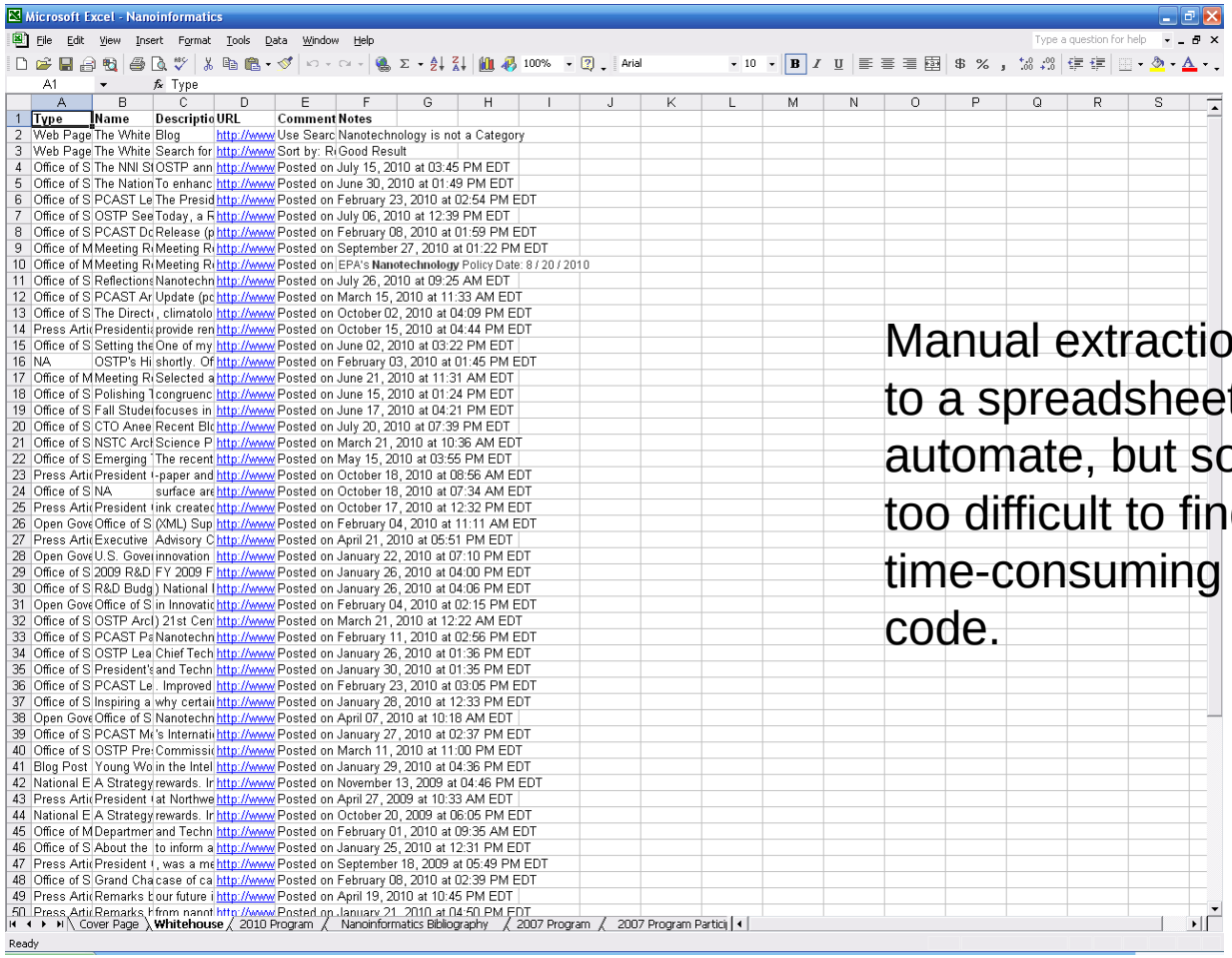
Need: A database to keep track of this and everything else in nanotechnology/ informatics!



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

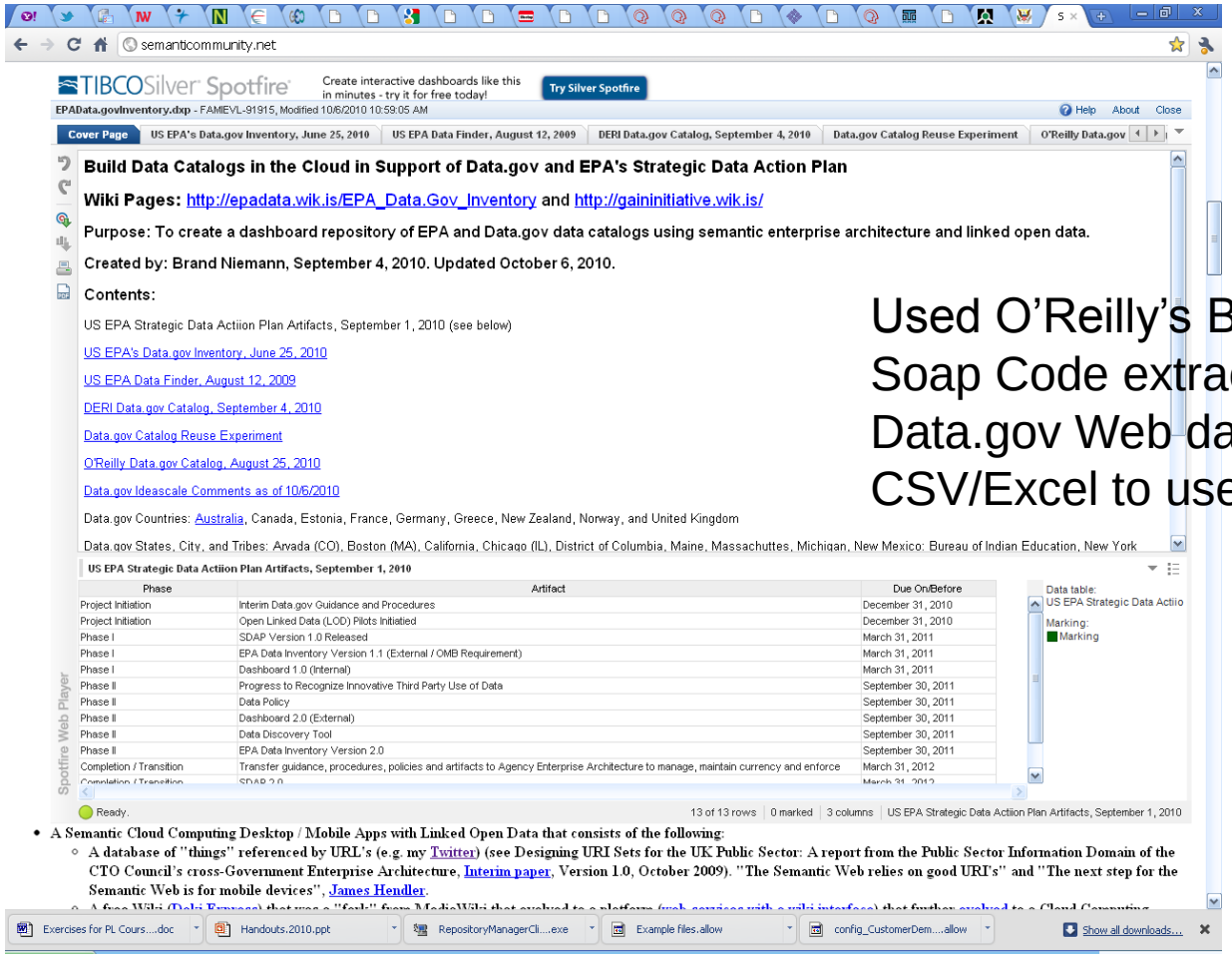
White House Blog: Nanotechnology



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



Used O'Reilly's Beautiful Soap Code extraction of the Data.gov Web database into CSV/Excel to use in Spotfire.

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

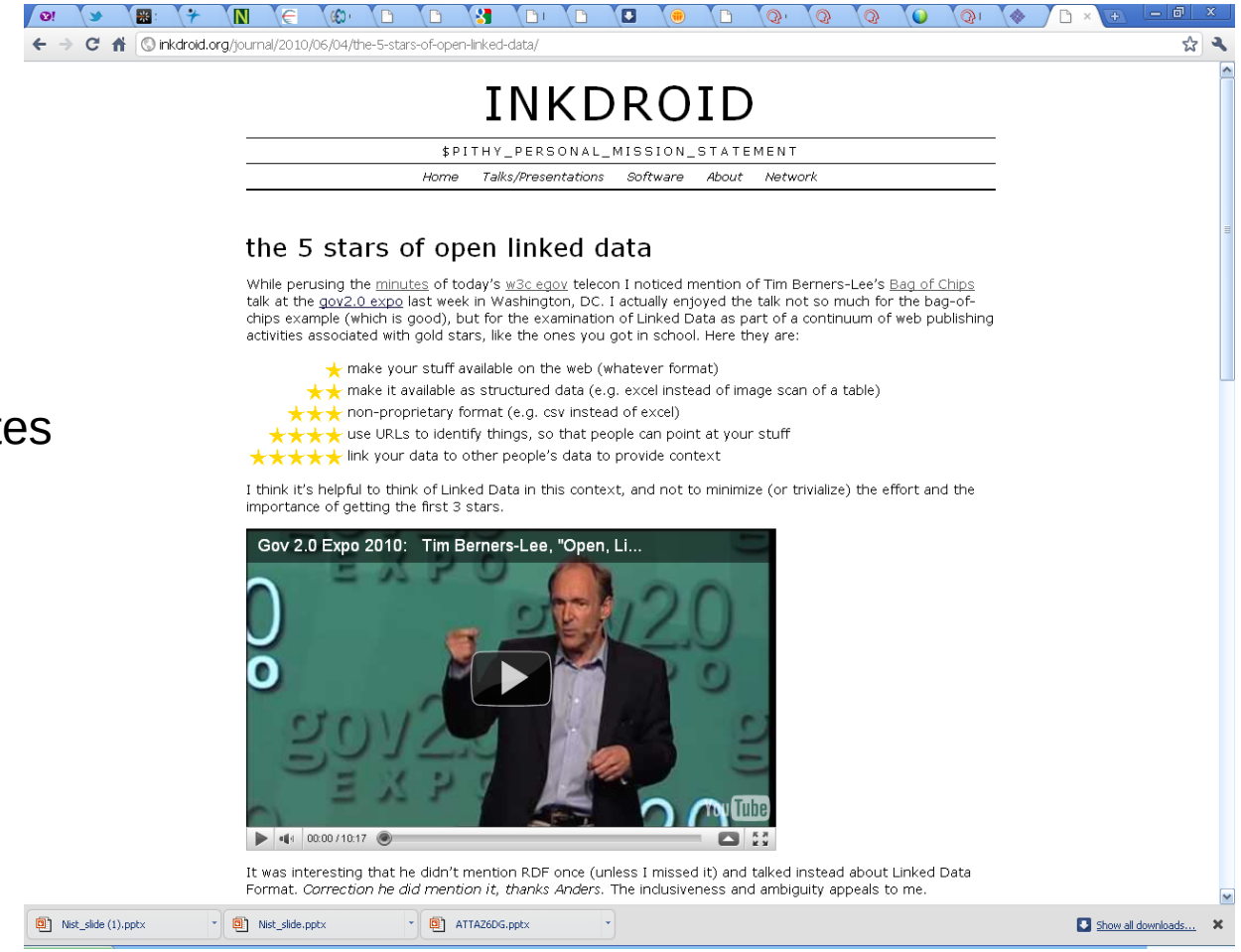
http://semanticommunity.wik.is/Nanoinformatics

Nanoinformatics 2007 in a Wiki

http://semanticcommunity.wik.is/Nanoinformatics/Participants/Mark_Tuominen

The Five Stars of Linked Open Data

YouTube:
10:17 Minutes

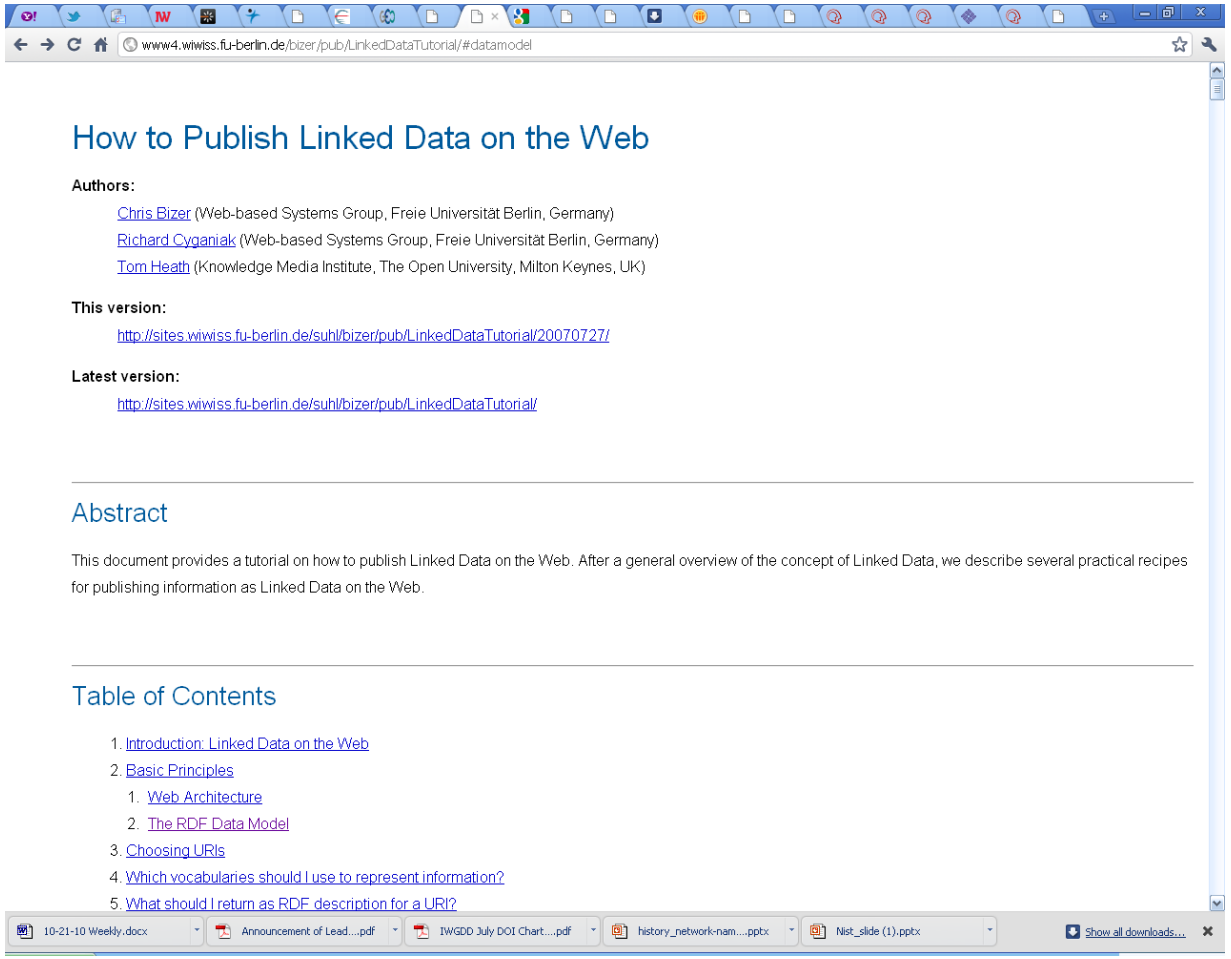


http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/

The Five Stars of Linked Open Data

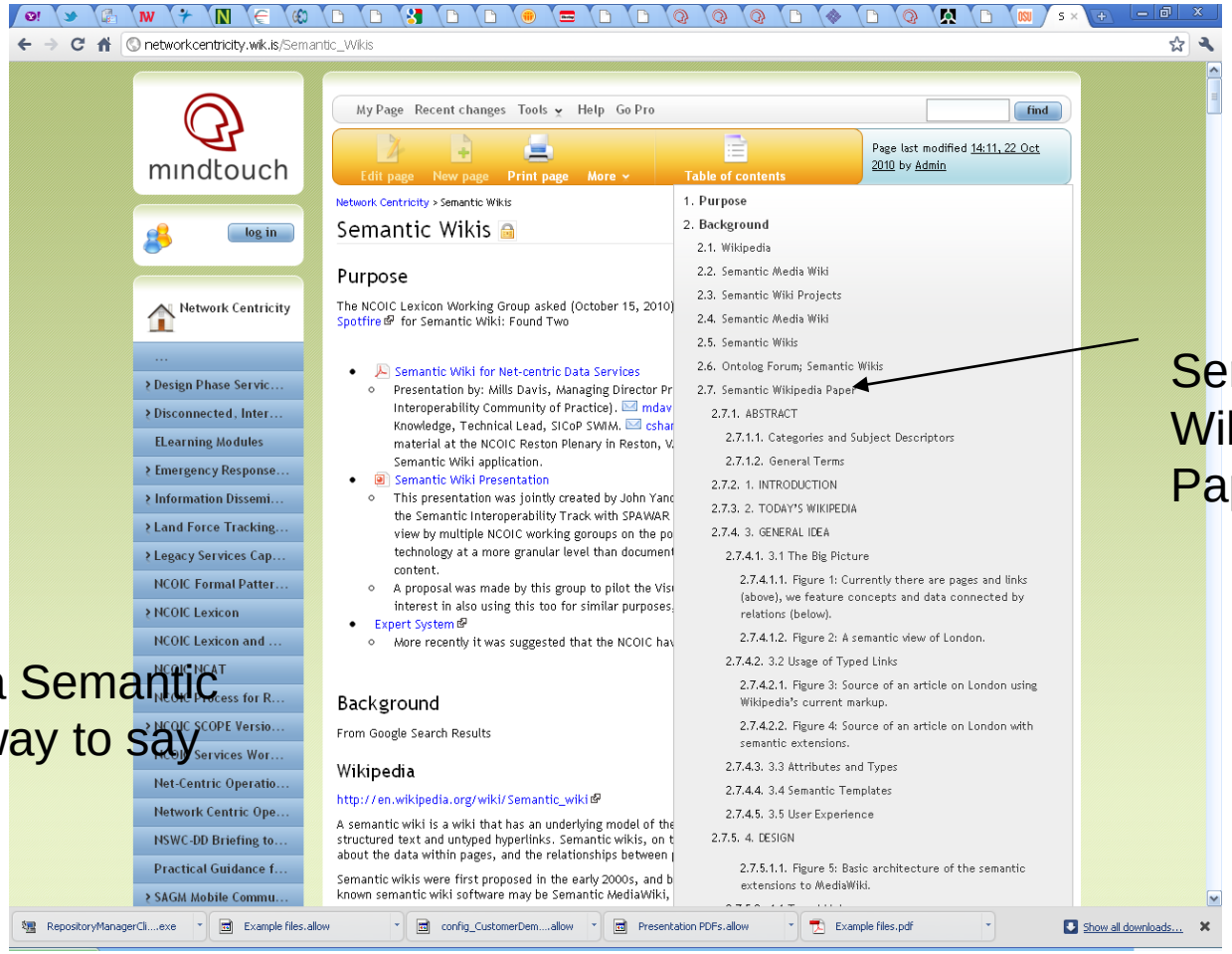
- 1. Make your stuff available on the web (whatever format);
- 2. Make it available as structured data (e.g. excel instead of image scan of a table);
- 3. Non-proprietary format (e.g. csv instead of excel);
- 4. Use URLs to identify things, so that people can point at your stuff; and
- 5. Link your data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

How to Publish Linked Data on the Web



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

Semantic Wikis

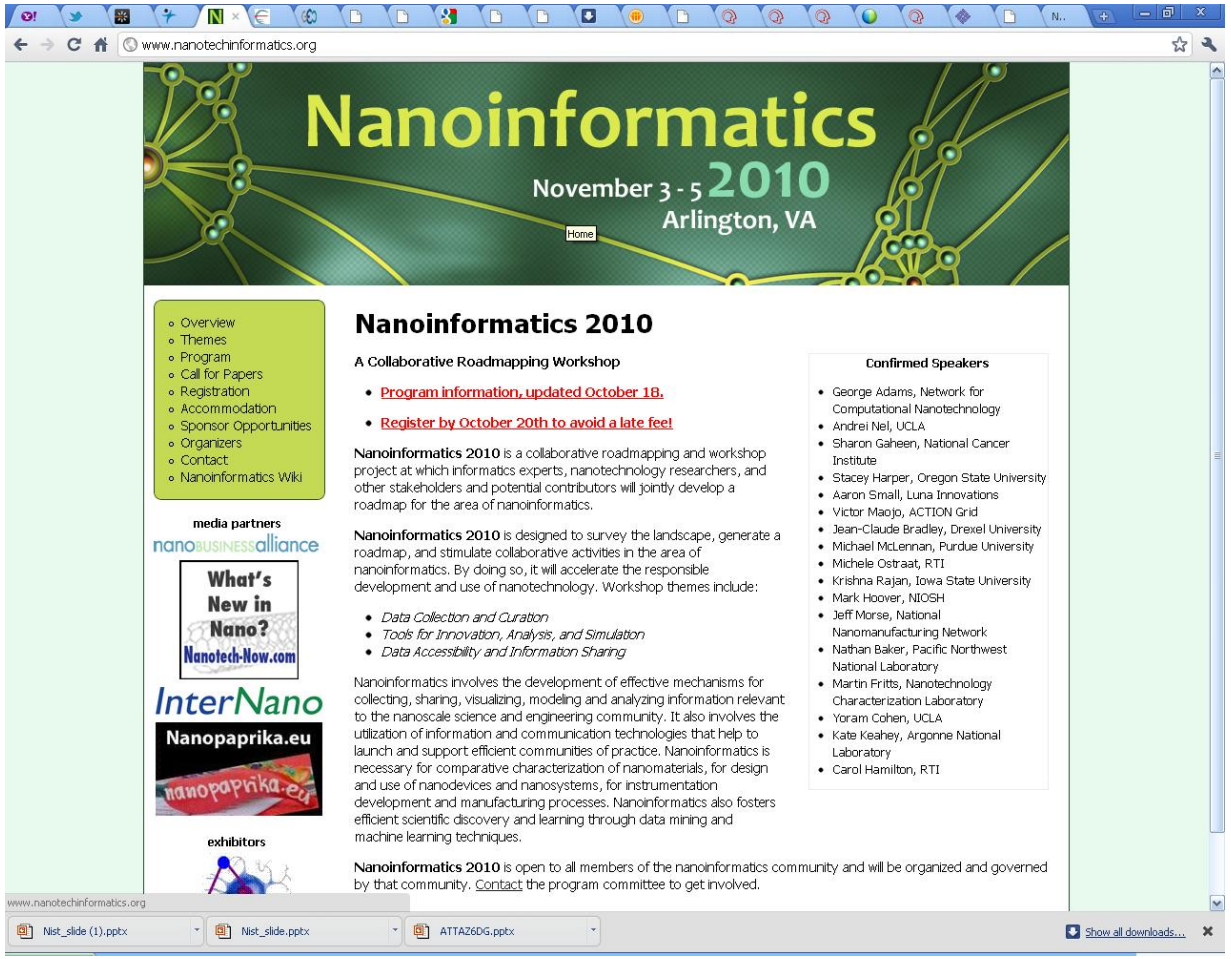


Semantic Wikipedia Paper (2006)

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

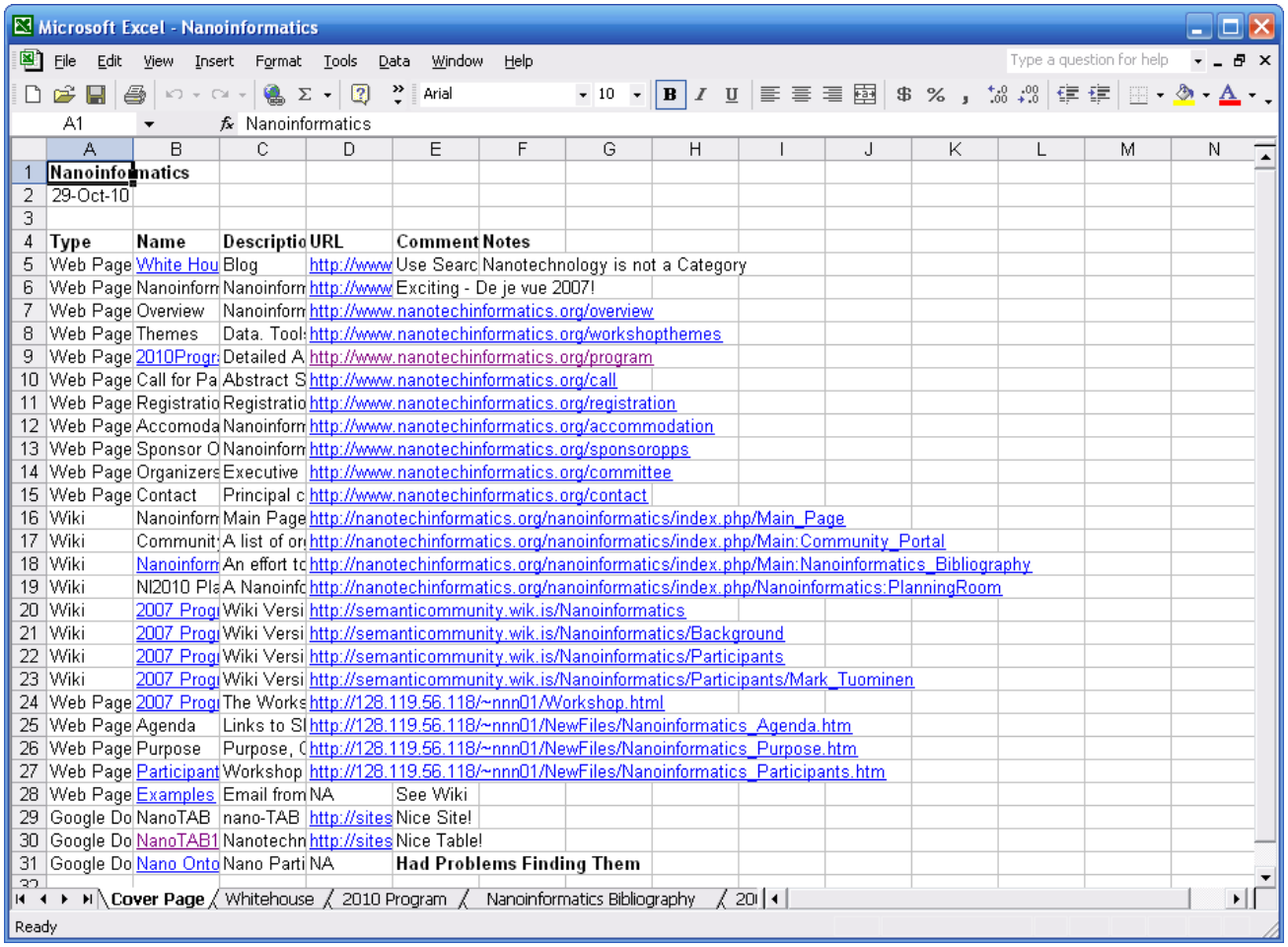
http://networkcentricity.wik.is/Semantic_Wikis

1. Make your stuff available on the web



http://www.nanotechinformatics.org/

2. Make it available as structured data



Microsoft Excel - Nanoinformatics

	A	B	C	D	E	F
1	Nanoinformatics					
2	29-Oct-10					
3						
4	**Type**	**Name**	**Descriptio**	**URL**	**Comment**	**Notes**
5	Web Page	White Hou	Blog	http://www	Use Searc	Nanotechnology is not a Category
6	Web Page	Nanoinform	Nanoinform	http://www	Exciting - De je vue 2007!	
7	Web Page	Overview	Nanoinform	http://www.nanotechinformatics.org/overview		
8	Web Page	Themes	Data. Tool	http://www.nanotechinformatics.org/workshopthemes		
9	Web Page	2010Progr	Detailed A	http://www.nanotechinformatics.org/program		
10	Web Page	Call for Pa	Abstract S	http://www.nanotechinformatics.org/call		
11	Web Page	Registratio	Registratio	http://www.nanotechinformatics.org/registration		
12	Web Page	Accomoda	Nanoinform	http://www.nanotechinformatics.org/accommodation		
13	Web Page	Sponsor O	Nanoinform	http://www.nanotechinformatics.org/sponsoropps		
14	Web Page	Organizers	Executive	http://www.nanotechinformatics.org/committee		
15	Web Page	Contact	Principal c	http://www.nanotechinformatics.org/contact		
16	Wiki	Nanoinform	Main Page	http://nanotechinformatics.org/nanoinformatics/index.php/Main_Page		
17	Wiki	Communit	A list of or	http://nanotechinformatics.org/nanoinformatics/index.php/Main:Community_Portal		
18	Wiki	Nanoinform	An effort to	http://nanotechinformatics.org/nanoinformatics/index.php/Main:Nanoinformatics_Bibliography		
19	Wiki	NI2010 Pla	A Nanoinfo	http://nanotechinformatics.org/nanoinformatics/index.php/Nanoinformatics:PlanningRoom		
20	Wiki	2007 Progr	Wiki Versi	http://semanticommunity.wik.is/Nanoinformatics		
21	Wiki	2007 Progr	Wiki Versi	http://semanticommunity.wik.is/Nanoinformatics/Background		
22	Wiki	2007 Progr	Wiki Versi	http://semanticommunity.wik.is/Nanoinformatics/Participants		
23	Wiki	2007 Progr	Wiki Versi	http://semanticommunity.wik.is/Nanoinformatics/Participants/Mark_Tuominen		
24	Web Page	2007 Progr	The Works	http://128.119.56.118/~nnn01/Workshop.html		
25	Web Page	Agenda	Links to SI	http://128.119.56.118/~nnn01/NewFiles/Nanoinformatics_Agenda.htm		
26	Web Page	Purpose	Purpose, (	http://128.119.56.118/~nnn01/NewFiles/Nanoinformatics_Purpose.htm		
27	Web Page	Participant	Workshop	http://128.119.56.118/~nnn01/NewFiles/Nanoinformatics_Participants.htm		
28	Web Page	Examples	Email from	NA	See Wiki	
29	Google Do	NanoTAB	nano-TAB	http://sites	Nice Site!	
30	Google Do	NanoTAB1	Nanotechn	http://sites	Nice Table!	
31	Google Do	Nano Onto	Nano Parti	NA	**Had Problems Finding Them**	

Cover Page | Whitehouse | 2010 Program | Nanoinformatics Bibliography | 20

Ready

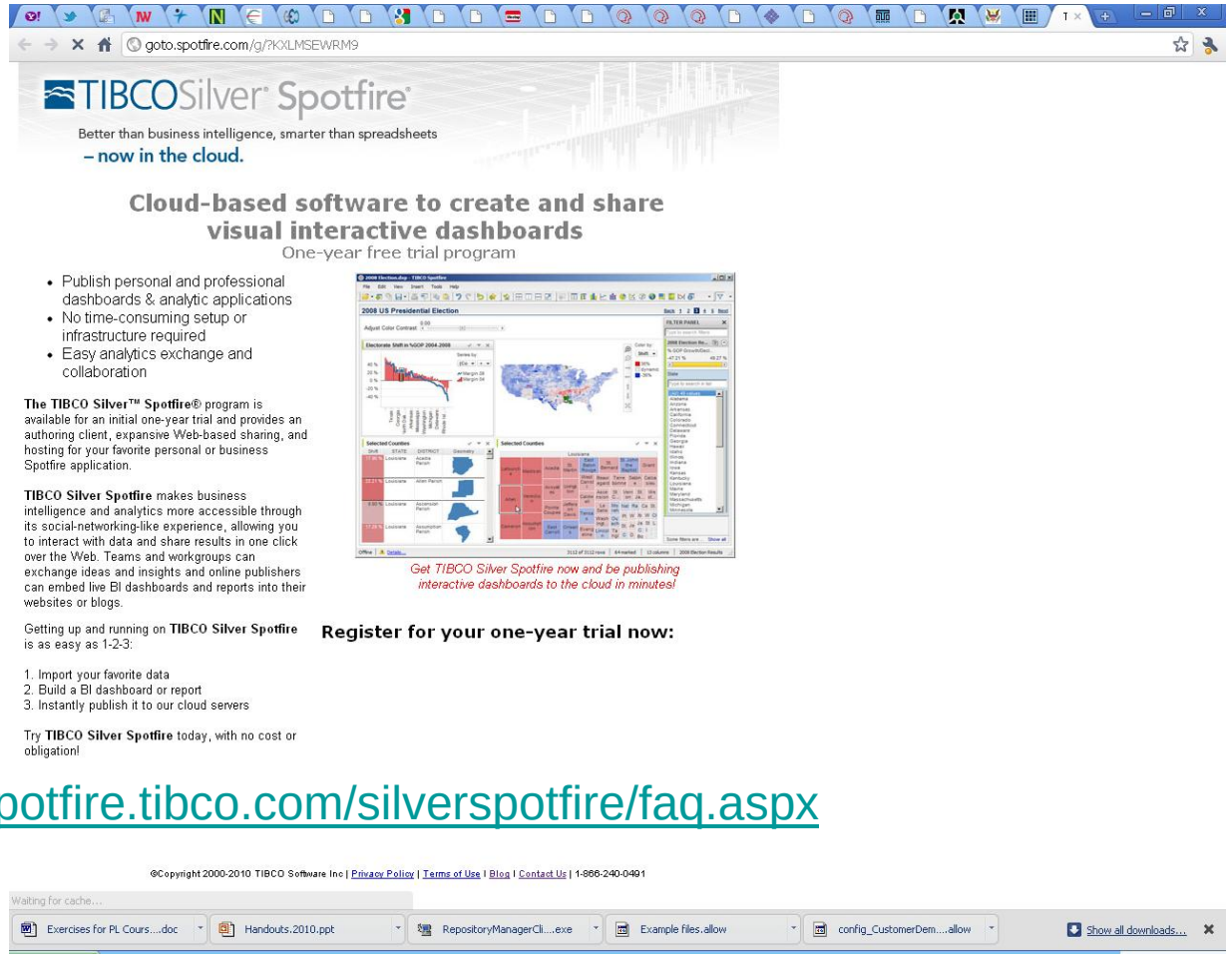
http://semanticommunity.wik.is/@api/deki/files/1750/=Nanoinformatics.xls

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

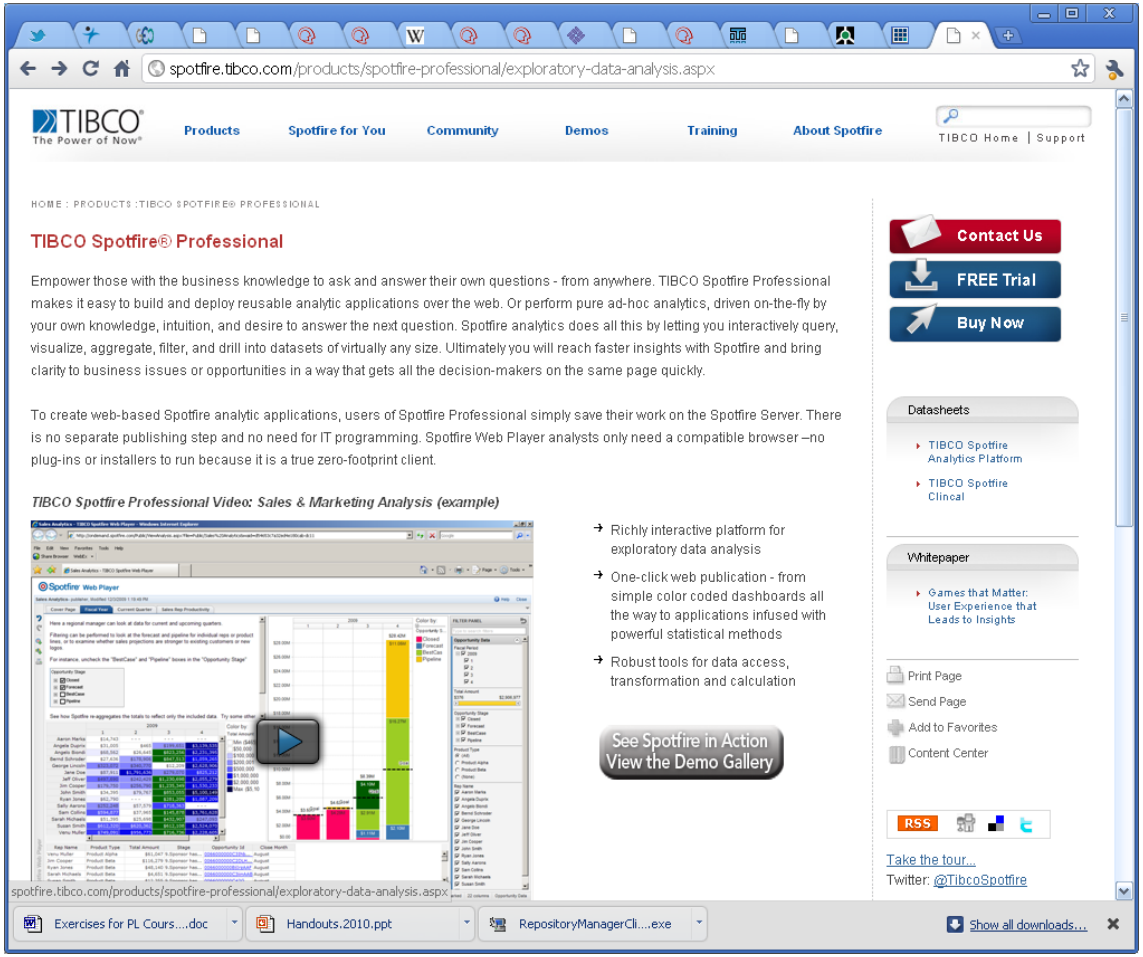


See: http://spotfire.tibco.com/silverspotfire/faq.aspx

http://goto.spotfire.com/g/?KXLMSEWRM9

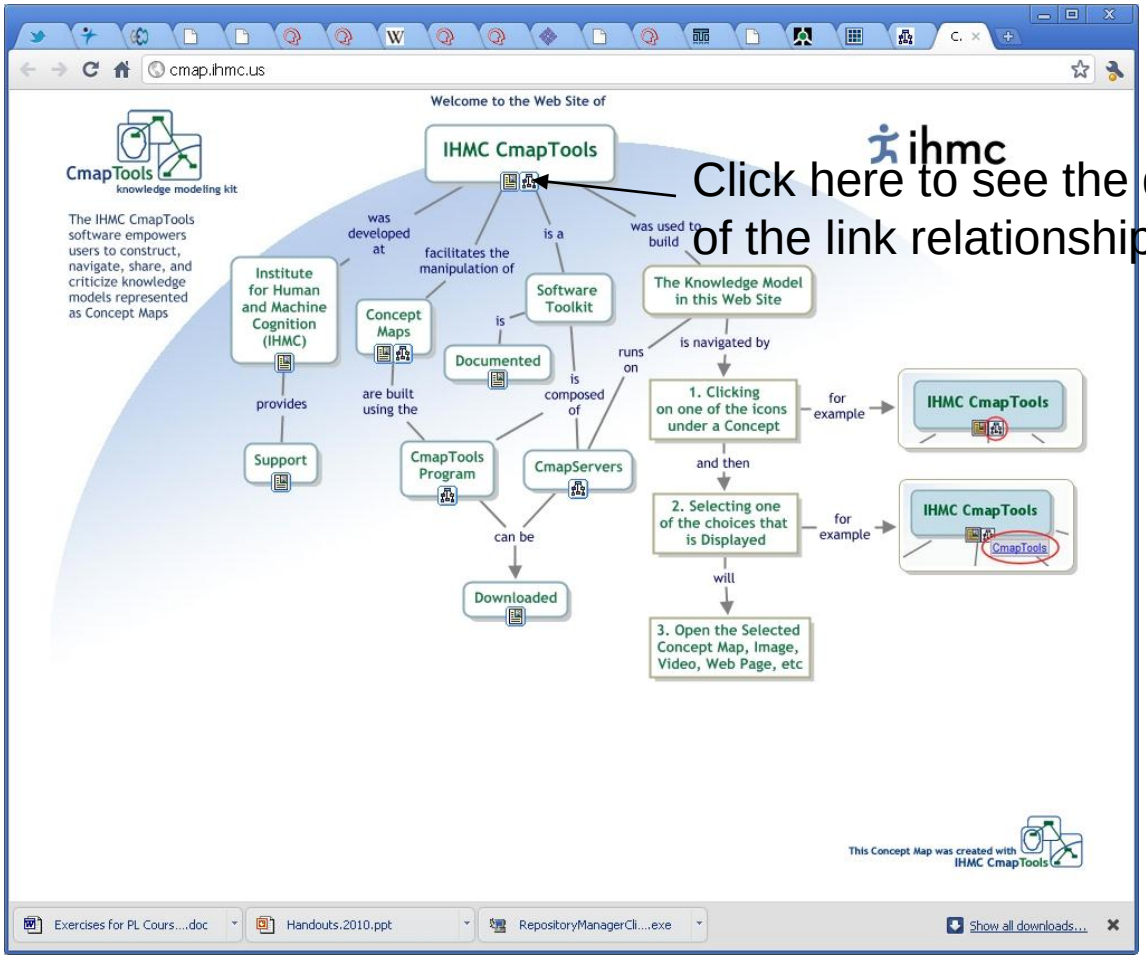
3A. TIBCO Spotfire Professional

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



http://spotfire.tibco.com/products/spotfire-professional/exploratory-data-analysis.aspx

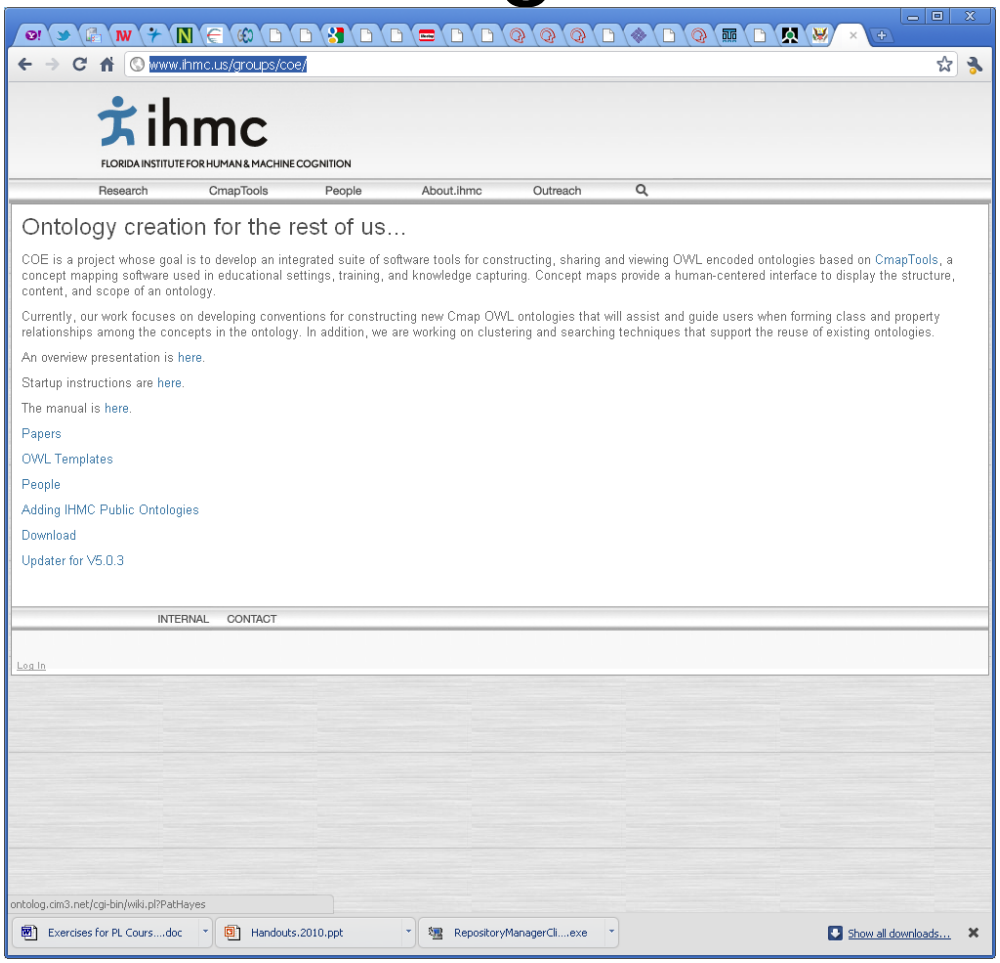
3B. IHMC Cmap Tools



Click here to see the details of the link relationships

http://cmap.ihmc.us/

3B. Florida Institute for Human and Machine Cognition COE



http://www.ihmc.us/groups/coe/

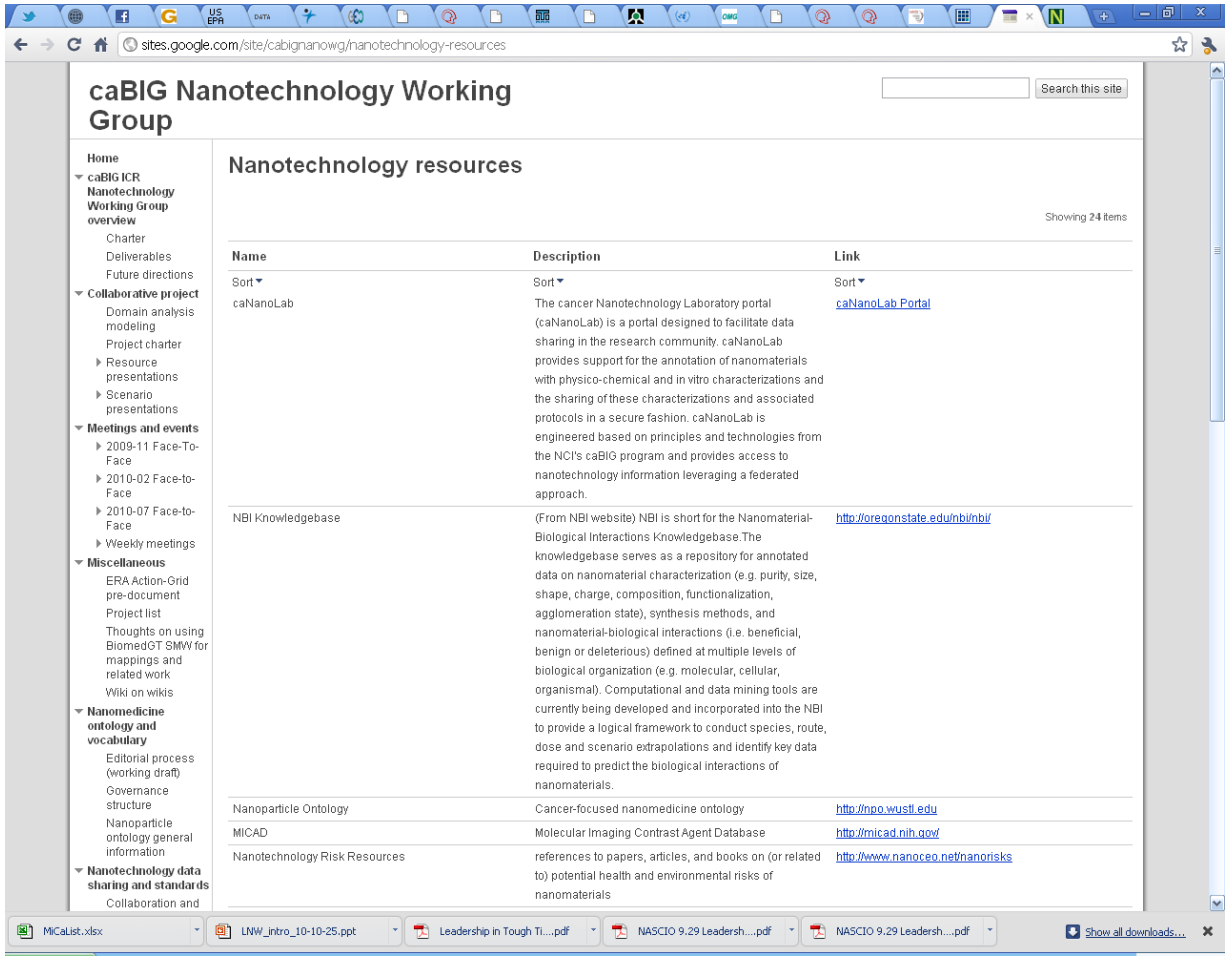
4. Use URLs to identify things, so that people can point at your stuff

- 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



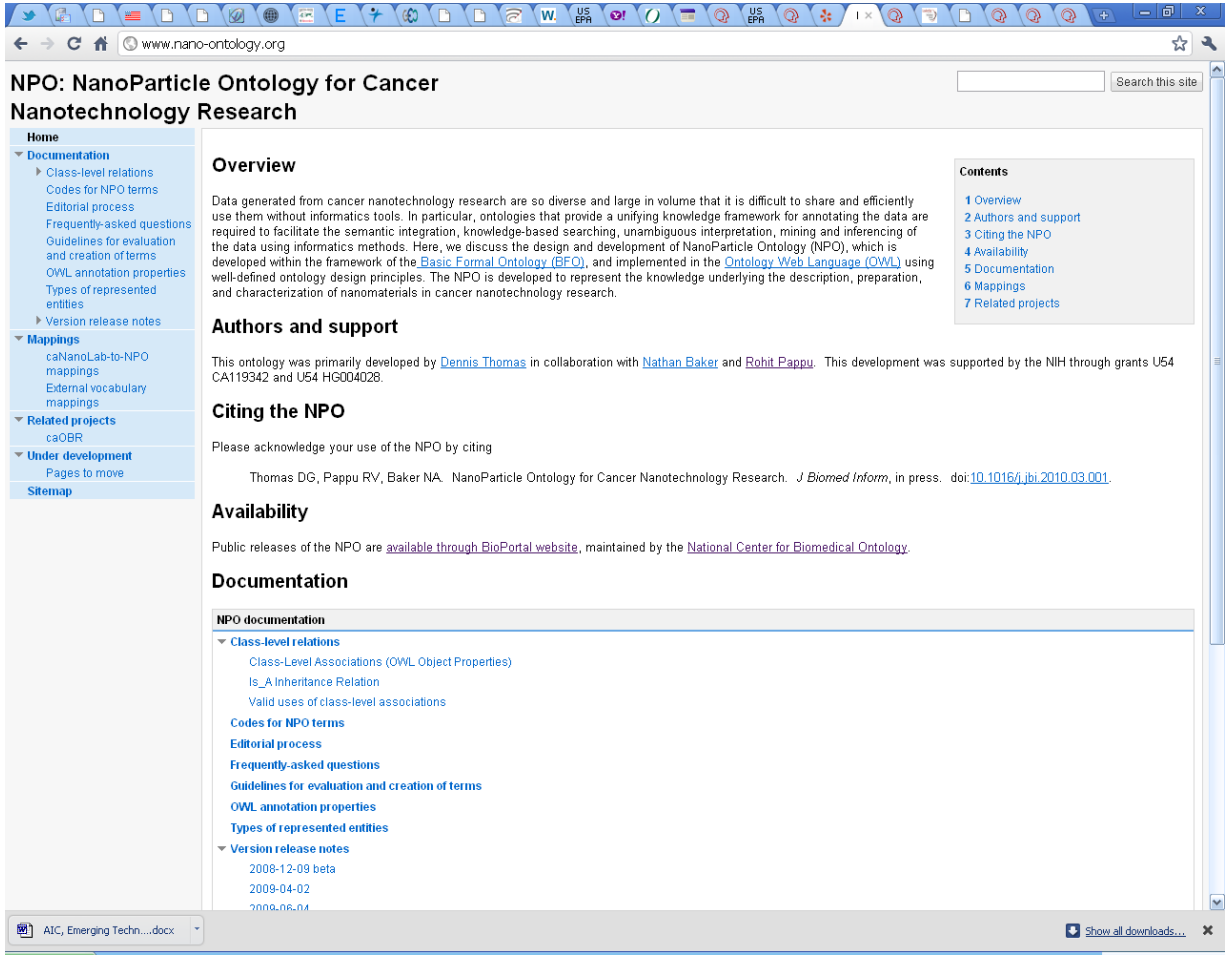
http://sites.google.com/site/cabignanowg/nanotechnology-resources

Nano Data in Excel

Download Excel File MiCaList.xls

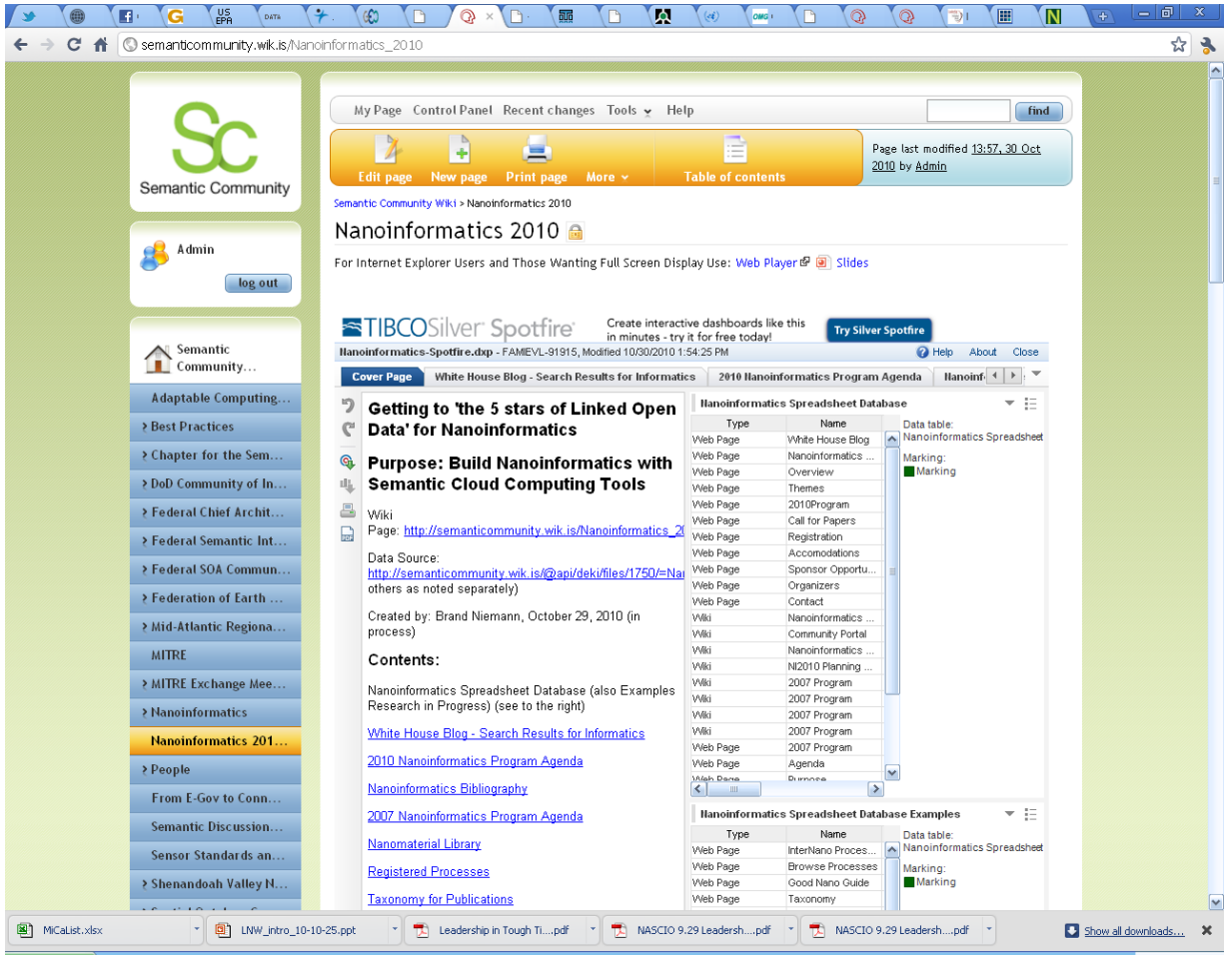
http://www.ncbi.nlm.nih.gov/bookshelf/corehtml/pmc/homepages/bookshelf/micad.html

NanoParticle Ontology for Cancer Nanotechnology Research



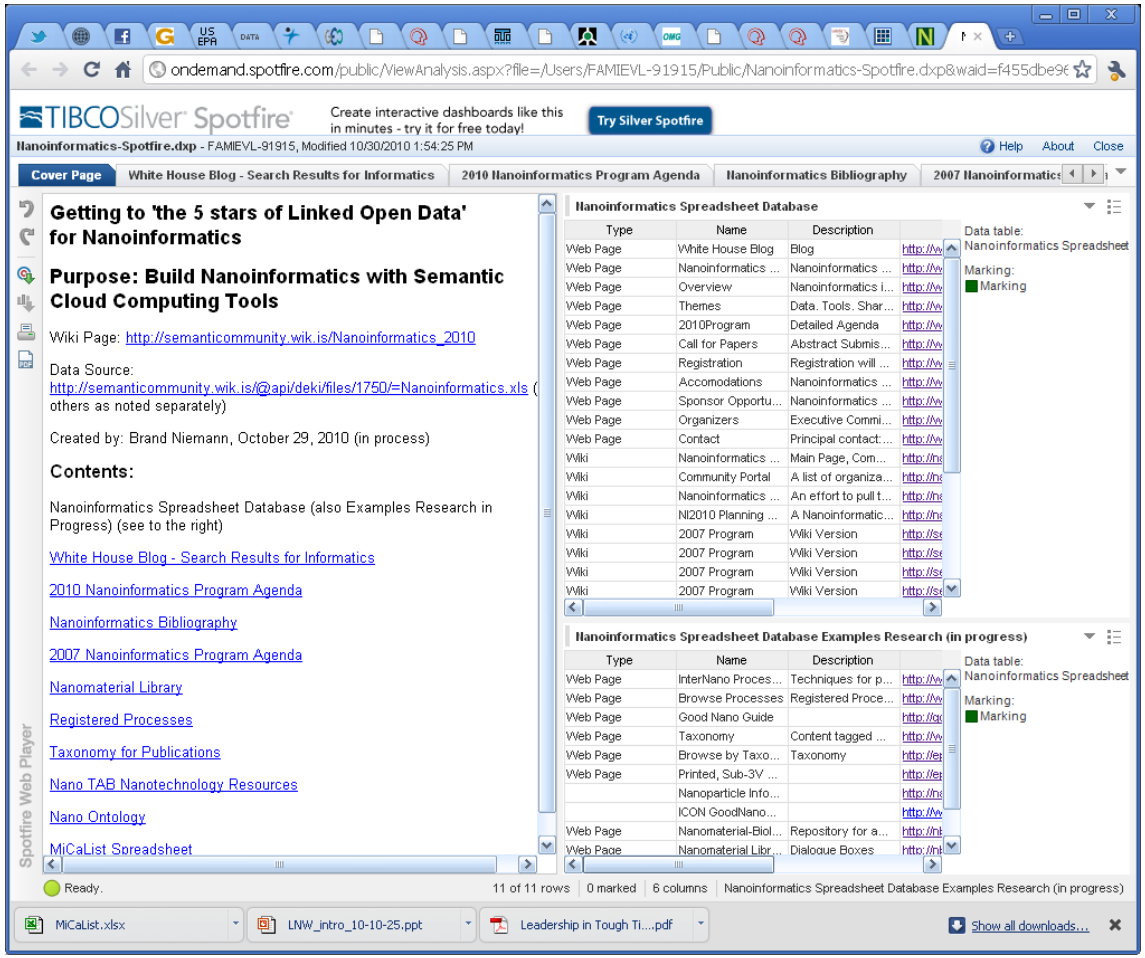
http://www.nano-ontology.org/

Build Nonoinformatics with Semantic Cloud Computing Tools



http://semanticommunity.wik.is/Nanoinformatics_2010

Build Nonoinformatics with Semantic Cloud Computing Tools



Spotfire Web Player

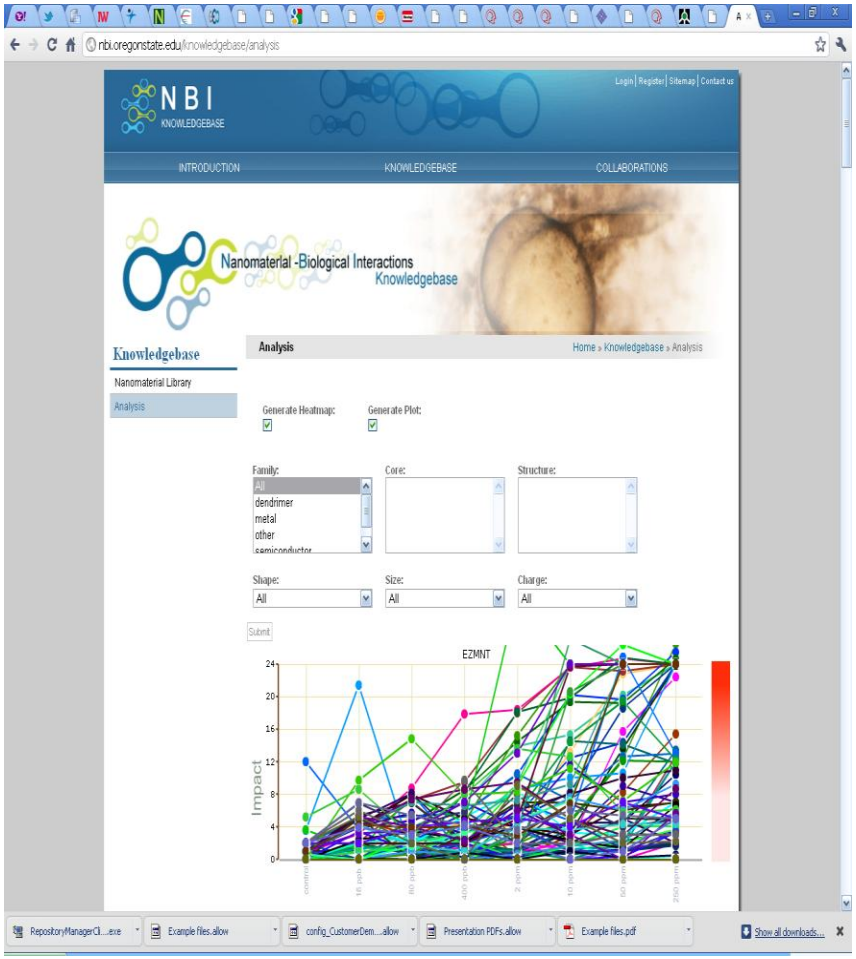
Some Next Steps

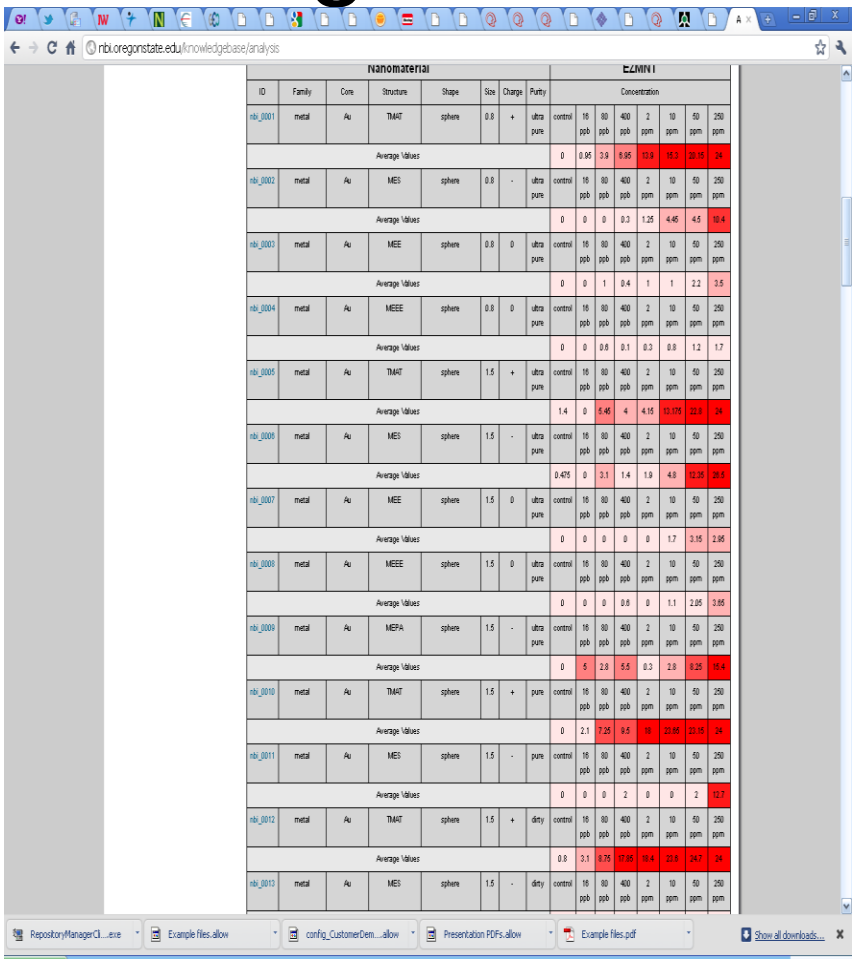
- Continue with some more manual extraction of key Nanoinformatics web pages to a spreadsheet.
- Use (or write) simple codes to automatically extract key Nanoinformatics databases to CSV.
- Inventory key Nanoinformatics databases in CSV (Excel) and import them into Spotfire for Linked Data Analytics.
- 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

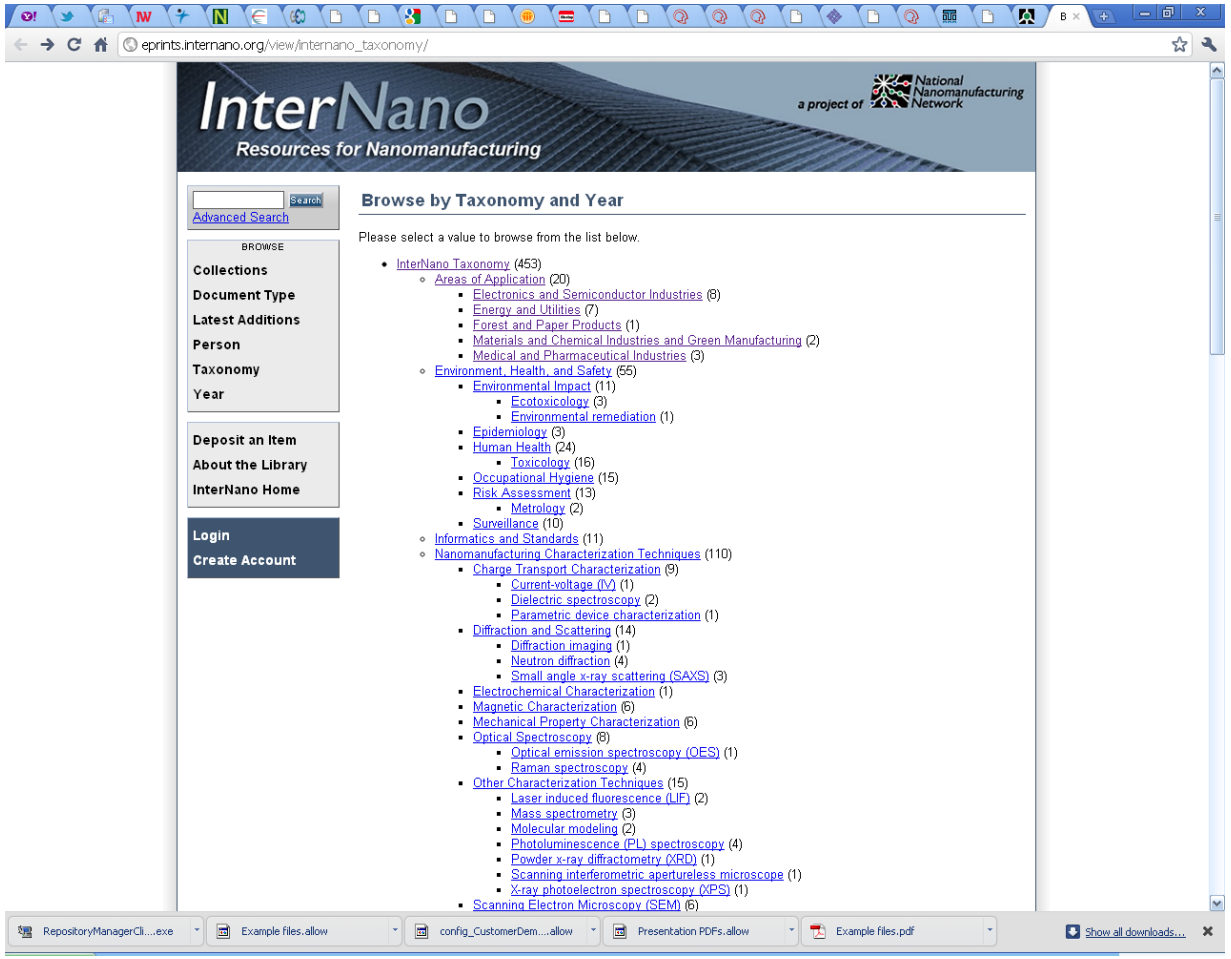
Nanomaterial-Biological Interactions Knowledgebase





http://nbi.oregonstate.edu/knowledgebase/analysis

InterNano Taxonomy: Top Level



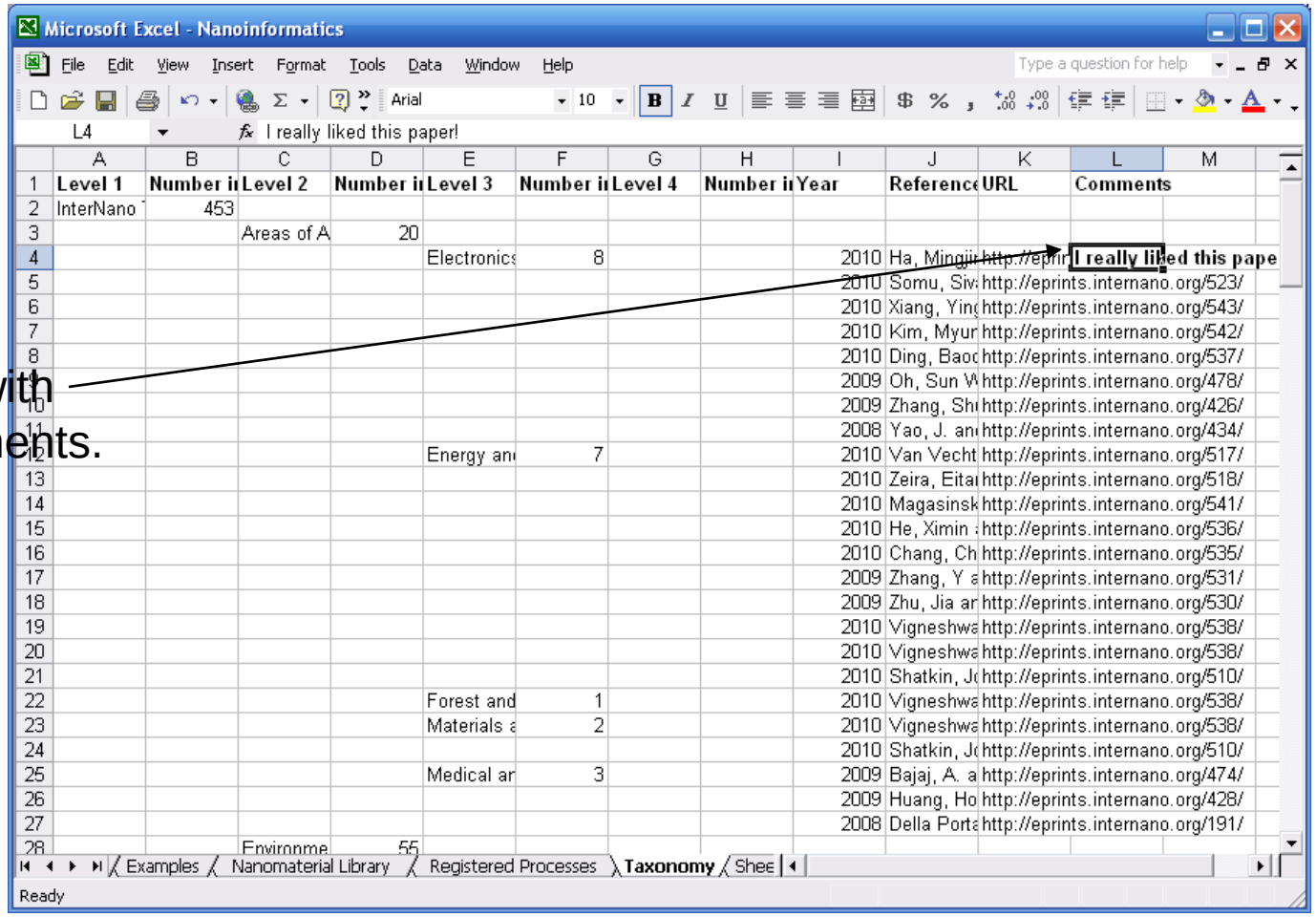
http://eprints.internano.org/view/internano_taxonomy/

InterNano Taxonomy: Bottom Level

http://eprints.internano.org/view/internano_taxonomy/internanotaxonomy5827/

IntreNano Taxonomy: Spreadsheet

Annotate with your comments.



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