Chemo::Blogs #2

Egon Willighagen ⓘ

Published December 6, 2006

Citation

Willighagen, E. (2006). Chemo::Blogs #2. In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/gwy1n-1sc04

Keywords

Blog, Rdf, Textmining, Cb

Abstract

Because no one picked up my Chemo::Blogs suggestion, I will now officially claim the blog series title. However, unlike the original Bio::Blogs series, I will not summarize interesting blogs, but just spam you with websites I recently marked as toblog on del.icio.us.

Copyright

Chemo::Blogs #2

chem-bla-ics

Because no one picked up my Chemo::Blogs suggestion, I will now officially claim the blog series title. However, unlike the original Bio::Blogs series, I will not summarize interesting blogs, but just spam you with websites I recently marked as toblog on del.icio.us.

Semantics and Text Mining

Evan Prodromou wrote about RDFa vs microformats. The latter are commonly used in enhancing blog semantics , and for example used by PostGenomic.com. While RDFa is more explicit, e.g. by using namespaced markup, we have to wait until XHTML2 to see it working. I do not think chemists are using tags a log yet, but let me propose the following microformats: 1/CH4/h1H4 and methane. Standard JavaScripts and CSS scripts will then do the rest. (Think: addressing newlines, auto googling-for-inchi, etc).

The reason why using microformats is interesting, is text mining, of various kinds. Whether it is setting up a molecule-article link database, or find hot molecules in blogspace , adding semantics will help tools like OSCAR3 to mine chemistry . Some time ago OTMI was proposed by Nature , and they now set up a dedicated web site to explain there view on text mining. Zack Rosen has a good idea why RDF Semantic web research isn't working.

Blogspace

There are a few new chemistry blogs I want to mention (and already added to Chemical blogspace): ChemBark, lirico which has an interesting chemoinformatics section, and The Curious Wavefunction. Worth reading indeed.

Pierre's YOKOFAKUN deserves a paragraph of his own. He recently blogged about bio2rdf which provides an RDF interface to biochemical knowledge via Life Science Identifiers (LSID), OBOEdit which is a Java-based ontology editor, and Amadea which is a Taverna- and KNIME-like tool for setting up UNIX pipes.

Online EMBL Symposium

A few EMBL PhD students are having the First Online EMBL PhD Symposium (catchy name, or … ;) Anyway, discussions are held on IRC, and it has a rather interesting Web2.0 session. All media is available on the website but requires registration right now. After the conference it will become open access to all. Jean-Claude contributed *The UsefulChem Project: Open Source Chemistry Research using Blogs and Wikis* to the Participants' Contributions section, and I had a poster on *Distributing molecular information over the Internet*, discussing CMLRSS, blog aggregators, CML and other things. The IRC session was logged and is available here.

Literature

Finally, I want to mention three recent articles. First one is a recent write up by Bourne and Friedberg about *Ten Simple Rules for Selecting a Postdoctoral Position* (DOI: 10.1371/journal.pcbi.0020121). With the end of my current postdoc position nearing, rather useful reading. Some time ago I blogged about a New open access journal Source Code for Biology and Medicine , and the journal is now up and running. Details can be read in the first editorial (DOI: 10.1186/1751-0473-1-1). The third article I would like to mention is *Scientific Software Development Is Not an Oxymoron* by Baxter (DOI: 10.1371/journal.pcbi.0020087), though I do not think it has new insights.

OK, this was a rather lengthy write up, but really needed to clean up my toblog section :)