chem-bla-ics

Molecules in Wikipedia

Egon Willighagen

Published August 2, 2007

Citation
Willighagen, E. (n.d.). In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/jv7te-h3w77

Keywords
Chemistry, Wikipedia, Rdf, Inchi

Copyright
Copyright © Egon Willighagen 2007. Distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

I do not care about physical and chemical properties in Wikipedia, as I can easily extract them from other sources. The main value of Wikipedia for molecules is, I think, that it describes the history of a molecule. Additionally, the Wikipedia URL is a nice unique molecular identifier (for example *http://en.wikipedia.org/wiki/Lactose*) given certain conditions, and many bloggers are using it as such . But, it only is a useful identifier if one (and only one) InChI is stated on the wiki page.

Now that I am RDF-ing molecular space , I was again interested in dbpedia, a RDF version of Wikipedia. See these two blog items and Peter's very nice dbpedia, RDF and SPARQL - for chemistry item. Christian is picking this up, and extending dbpedia for support for the various chemical boxes.

Wikipedia Templates

I have spotted a couple of templates: Drugbox, Chembox, Chembox new, of which the last one seems to most recent, and has extensions for explosives and drugs. The WikiProject Chemicals does not mention it though. Anyone who knows the status? Is *chembox new* the way forward and going to replace the older *chembox*? I hope so, because only the newer one has InChI in the last of official fields. Or is *chembox new* simply an extension of *chembox* itself?

Somewhere between 1000 and 1500 entries use the *chembox new* and another 1000 to 1500 use *chembox* but I assume there is considerable overlap. Additionally, Christian noted that there still seem to be molecules in Wikipedia which do not use a template at all, and counted some 1900 molecules using various lists. If you you want to keep a more close eye on chemistry in dbpedia, you should register to the dbpedia-discussion mailing list.