

DBPedia: lookup and autocomplete of chemistry

Egon Willighagen 

Published February 11, 2009

Citation

Willighagen, E. (2009). DBPedia: lookup and autocomplete of chemistry. In *chem-bla-ics*. chem-bla-ics. <https://doi.org/10.59350/wbb63-gv107>

Keywords

Rdf, Dbpedia, Wikipedia, Smiles, Inchi

Copyright

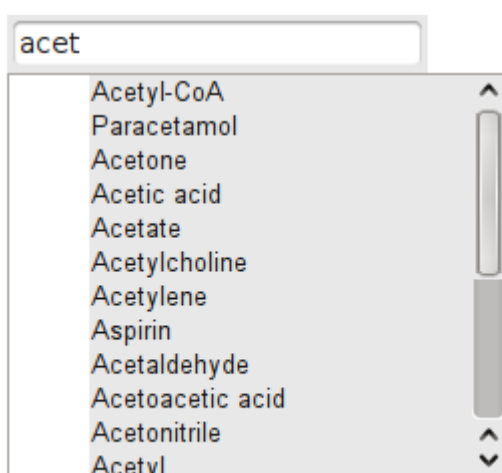
Copyright © Egon Willighagen 2009. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

On the [DBpedia discussion mailing list](#) there was a post on a nice web page which allows you to look up things, and which features an autocomplete edit field. The below screenshot shows lookup of molecular structures:



URI Lookup Autocomplete Test



If you are not aware of this, adding content to DBpedia is as easy as adding something to [Wikipedia](#). Literally: DBpedia is the [RDF](#) flavour of Wikipedia. It extracts the information from the info boxes, as I discussed before (see [Molecules in Wikipedia](#)).

BTW, one can take advantage of DBpedia to see what Wikipedia has to offer in terms of chemistry. For example, to list all molecules which have a SMILES, one can use this simple SPARQL query:

Or, to list those which have an InChI:

And this is actually quite useful, e.g. it can be used in quality control. Running the above queries will show up several broken SMILES and InChIs. I have not had time to fix those yet, so please go ahead and beat me to those fixes, and get some Wikipedia Fame :) Alternatively, invert the queries and add missing InChIs, PubChem CID or SMILES. When I have a bit more free time again, after the new stable CDK and Bioclipse releases, I'll run these analyses again, and summarize them in a web page.