BioSpider: another molecule search engine

Egon Willighagen

Published October 29, 2007

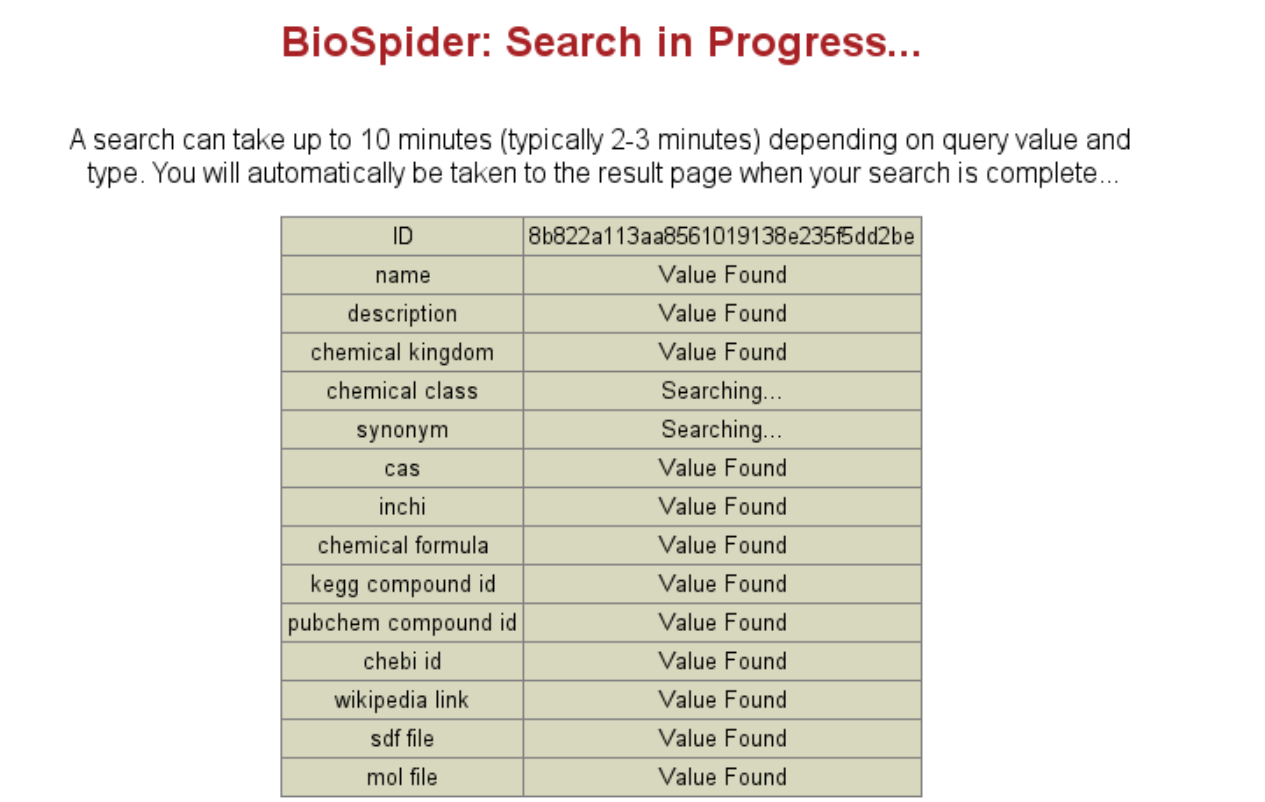
Citation

Willighagen, E. (2007). BioSpider: another molecule search engine. In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/9qsxx-j6z92

Abstract

I just ran into BioSpider. Unlike ChemSpider, BioSpider crawls the internet (well, this list of sources really) to find information, and depending on what it finds it continues the search.

I just ran into BioSpider. Unlike ChemSpider, BioSpider crawls the internet (well, this list of sources really) to find information, and depending on what it finds it continues the search. Below is a screenshot of an intermediate point after starting with the InChI of methane:



BioSpider: Search in Progress...

A search can take up to 10 minutes (typically 2-3 minutes) depending on query value and type. You will automatically be taken to the result page when your search is complete...

ID	8b822a113aa8561019138e235f5dd2be
name	Value Found
description	Value Found
chemical kingdom	Value Found
chemical class	Searching...
synonym	Searching...
cas	Value Found
inchi	Value Found
chemical formula	Value Found
kegg compound id	Value Found
pubchem compound id	Value Found
chebi id	Value Found
wikipedia link	Value Found
sdf file	Value Found
mol file	Value Found

After the search it generates a long HTML page with all the information it found on the molecule you queried for. This approach is much more scalable than storing all in one database.

This crawling of information is something I was working on myself a bit too, and I think this is a good approach. However, I think the use of a central website is not the right approach. Instead, the search should be distributed too: the crawling should be done on the client machine; it should be done in Taverna or Bioclipse instead.

My conclusion: excellent idea, bad implementation.