chem-bla-ics

Solubility Data in Bioclipse #3: Finding ChEBI IDs

Egon Willighagen

Published February 27, 2009

Citation
Willighagen, E. (2009, February 27). Solubility Data in Bioclipse #3: Finding ChEBI IDs. *Chem-bla-ics*. https://doi.org/10.59350/gh3np-xbm68

Keywords
Gist, Sparql, Rdf, Chebi

Abstract
With the RDF functionality set up in Bioclipse (see Solubility Data in Bioclipse #2: handling RDF), we can start mining the Chemical RDF space.

chem-bla-ics

With the RDF functionality set up in Bioclipse (see Solubility Data in Bioclipse #2: handling RDF), we can start mining the Chemical RDF space. Check out this mashup:

What happens in this script is the following:

1. Load the ONS Solubility data (line 4-5)
2. ask for all owl:sameAs relations to navigate (line 8-14)
3. load the RDF for the rdf.openmolecule.net resources (line 16-26)
4. query for all solvents which have an ChEBI identifier (line 28-38)

The output will look like the following (in the future this will be opened as spreadsheet in Bioclipse):

```
[[ethanol 40C, CHEBI:16236],
[acetonitrile, CHEBI:38472],
[chloroform, CHEBI:35255],
[methanol 30C, CHEBI:17790],
[THF, CHEBI:26911],
[ethanol, CHEBI:16236],
[ethanol 30C, CHEBI:16236],
[methanol 40C, CHEBI:17790],
[methanol, CHEBI:17790]]
```

Now, this example shows a simple yet powerful feature of how RDF is used nowadays: the ChEBI identifier was not part of the original Solubility spreadsheet at Google Docs. But, taking advantage of the unique and *resolvable* URIs for molecules, when can simply look them up.

Nice, isn't it?