chem-bla-ics

Oscar text mining in Taverna

Published October 21, 2010

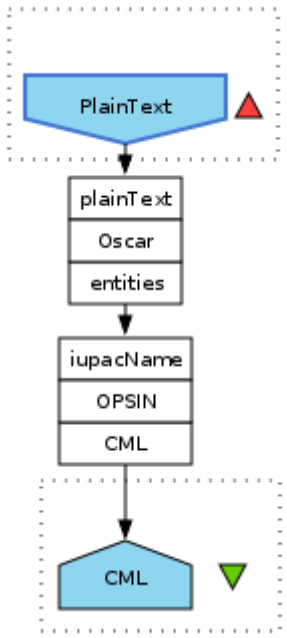
Citation
Willighagen, E. (2010, October 21). Oscar text mining in Taverna. *Chem-bla-ics*. https://doi.org/10.59350/7njvw-s6q24

Keywords
Oscar, Taverna

Abstract
One of the goals of my project in Cambridge is to make Oscar available as Taverna plugin (source code, Hudson build). I have progressed somewhat, but still struggling with getting the update site working.

chem-bla-ics

One of the goals of my project in Cambridge is to make Oscar available as Taverna plugin (source code, Hudson build). I have progressed somewhat, but still struggling with getting the update site working. The plugin actually installs into Taverna 2.2.0, but the activities do not show up. While this is work in progress, and the other project goal is refactoring, a current demo workflow looks like:



Example input would be: *This is a list of ethanol, methanol, and 2,4,6-trinitrotoluene.*

The plain text input can be linked to the pdf2text SADI service, and the CML is suitable for the CDK-Taverna plugin, which is currently being updated by Andreas, Achim, and Christoph for Taverna 2.2. As soon as the update site is properly working, I will upload a demo workflow to MyExperiment.org.

I guess the first next activity (node in the workflow) will be around the dictionaries, as the OPSIN activity converts only IUPAC names into connection tables. I was told OPSIN parses 97% of the IUPAC names it finds, and when it does, it does almost 100% correct. Want to challenge the code? Use this web service.