

Status update on BIOC analysis with Oscar and ChemicalTagger #3

Egon Willighagen 

Published December 23, 2010

Citation

Willighagen, E. (2010). Status update on BIOC analysis with Oscar and ChemicalTagger #3. In *chem-bla-ics*. chem-bla-ics. <https://doi.org/10.59350/8n7nt-fas57>

Keywords

Oscar, Chemicaltagger, Beilstein

Abstract

The two earlier posts in this series showed screenshots of results of Oscar, but the title also promised results by Lezan's ChemicalTagger. Sam helped with getting the HTML pages online via the Cambridge Hudson installation. Where Oscar finds named entities (chemical compounds, processes, etc), ChemicalTagger finds roles, like solvent, acid, base, catalyst. Roles are properties of chemical compounds in certain situations.

Copyright

Copyright © Egon Willighagen 2010. 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.

The two earlier posts in this series showed screenshots of results of Oscar, but the title also promised results by Lezan's [ChemicalTagger](#). Sam helped with getting the HTML pages online via the Cambridge Hudson installation. Where Oscar finds named entities (chemical compounds, processes, etc), ChemicalTagger finds roles, like solvent, acid, base, catalyst. Roles are properties of chemical compounds in certain situations. Ethanol is not always a solvent, sometimes it is a Xmas present. The current output is not entirely where I want to go yet, but makes it easy which solvents are frequently found in the BIOC corpus:

Compound	Role	Paper	Sentence
AcOEt	Solvent	PCM1399459	3-Acetyl-5-nitrohexanoic acid methyl ester (2a) Purified
AcOEt	Solvent	PCM1399459	3-Acetyl-5-oxohexanoic acid methyl ester (3a) Purified
AcOEt	Solvent	PCM1399459	4-Benzyl-3-methylcyclopent-2-enone (4d) Purified by c
brine	Solvent	PCM1399459	The organic layer was washed with brine and dried (M)
brine	Solvent	PCM2605621	The organic layer was washed with brine , dried over N
brine	Solvent	PCM2887305	The solution was washed successively with aqueous H- vacuo .
CDCl3	Solvent	PCM1399459	1H-NMR spectra (300 MHz) and 13C NMR spectra (75
cyclohexane	Solvent	PCM2605621	The UV / VIS-spectra of 2,3,6,7-tetrabromoanthracene a
cyclohexane	Solvent	PCM2605621	Figure 1 UV-spectra of 4 and commercially available anti
cyclohexane	Solvent	PCM2605621	Figure 1 UV-spectra of 4 and commercially available anti
dichloromethane	Solvent	PCM1399459	The mixture was treated with H2O , the methanol was e
dichloromethane	Solvent	PCM2956567	Ethyl 2-(4-hydroxy-2-oxabutyl)acrylate (2a) To a solution triethylamine in 200 mL dichloromethane was added slow
dichloromethane	Solvent	PCM2956567	The aqueous layer was reduced in volume to 100 mL an
diethyl ether	Solvent	PCM2887305	The resulting solid was collected by filtration through a 7 bottomed flask transferred to a round-bottomed flask dri
diethyl ether	Solvent	PCM2956567	Ethyl 2-(4-phosphonoxy-2-oxabutyl)acrylate (3a) To a 0.1008 mol) triethylamine in 250 mL diethyl ether was ac
diethyl ether	Solvent	PCM2956567	The layers were separated and the aqueous layer was

This screenshot of an analysis of 15 BIOC papers shows that AcOEt (is that the [same as EtOAc?](#)) is mentioned as solvent three times in [PCM1399459](#). Brine, however, is mentioned as solvent in three papers.

As said, these [two pages](#) contain RDF and the tables are sortable. Hudson recompiles them automatically when I update the source code to create the HTML+RDFa. So, go ahead, send me bug reports, feature requests, and patches!