

Resource Summary Report

Generated by [SPARC SAWG](#) on Aug 8, 2026

Oscar3

RRID:SCR_004561

Type: Tool

Proper Citation

Oscar3 (RRID:SCR_004561)

Resource Information

URL: <http://www-pmr.ch.cam.ac.uk/wiki/Oscar3>

Proper Citation: Oscar3 (RRID:SCR_004561)

Description: OSCAR is software for the semantic annotation of chemistry papers. The modules OPSIN (a name to structure converter) and ChemTok (a tokeniser for chemical text) are also available as standalone libraries. This tool for shallow, chemistry-specific parsing of chemical documents identifies (or attempts to identify): * Chemical names: singular nouns, plurals, verbs etc., also formulae and acronyms, some enzymes and reaction names. * Ontology terms: if you can do it by string-matching, you can get OSCAR to do it. * Chemical data: Spectra, melting/boiling point, yield etc. in experimental sections. In addition, where possible the chemical names that are detected are annotated with structures, either via lookup or name-to-structure parsing (OPSIN), and with identifiers from the chemical ontology ChEBI. Current work on OSCAR3 by Peter Corbett focuses on its use in SciBorg, a framework for the deep parsing of chemical text. OSCAR3 also includes the Oscar Server, a Jetty-powered set of servlets. These provide the following services: * Parsing of text/HTML by OSCAR. * Text/InChI/SMILES/SMILES substructures/SMILES similarity search of papers, coupled with keyword and ontology-based search, using Lucene and the CDK. * List of all names found / all names that co-occur with a search term or terms. * Online management of a chemical/stopword lexicon. * Manual editing of SciXML fragments containing named entities, for creating of gold standards and training data. Oscar3 can be found on SourceForge: <http://sourceforge.net/projects/oscar3-chem/>

Abbreviations: Oscar

Synonyms: OSCAR, Open Source Chemistry Analysis Routines

Resource Type: software resource

Defining Citation: [PMID:21999457](#)

Keywords: annotation, chemistry

Resource Name: Oscar3

Resource ID: SCR_004561

Alternate IDs: nlx_55584

Record Creation Time: 20220129T080225+0000

Record Last Update: 20260808T185823+0000

Ratings and Alerts

No rating or validation information has been found for Oscar3.

No alerts have been found for Oscar3.

Data and Source Information

Source: [SciCrunch Registry](#)

Usage and Citation Metrics

We found 3 mentions in open access literature.

Listed below are recent publications. The full list is available at [SPARC SAWG](#) .

Selcuk Y, et al. (2025) InfectA-Chat, an Arabic Large Language Model for Infectious Diseases: Comparative Analysis. JMIR medical informatics, 13, e63881.

Davis AP, et al. (2013) Text mining effectively scores and ranks the literature for improving chemical-gene-disease curation at the comparative toxicogenomics database. PloS one, 8(4), e58201.

Kolluru B, et al. (2011) Using workflows to explore and optimise named entity recognition for chemistry. PloS one, 6(5), e20181.