

Poster presentation

Free chemical structure search in MDL SDFiles using the open source Personal Chemistry Client (PCC)

Hans-Jürgen Himmler* and Alexander Kos

Address: AKos Consulting & Solutions Deutschland GmbH (AKos GmbH), Austr. 26, D-79585 Steinen, Germany

* Corresponding author

from 3rd German Conference on Chemoinformatics
Goslar, Germany. 11-13 November 2007

Published: 26 March 2008

Chemistry Central Journal 2008, 2(Suppl 1):P43 doi:10.1186/1752-153X-2-S1-P43

This abstract is available from: <http://www.journal.chemistrycentral.com/content/2/S1/P43>

© 2008 Himmler and Kos

The open source program Personal Chemistry Client (PCC) a .Net program integrates chemical structure information with associated alphanumeric data.

The user can choose to perform exact structure, substructure searches or structure similarity searches. The associated data can be stored in Microsoft Excel, and all major RDBMS systems (Oracle, Microsoft SQLServer and MySQL), as well as Microsoft Access and Comma separated files (CSV).

Out of-the-box the user can load MDL SDFiles and can search, browse and print the content of SDFiles. After the import the user can either load existing search forms or can define her/his own search form using simple drag and drop. After the search form is defined the user can paste structures in MDL Molfile format using e.g. ISIS/Draw, MDL Draw, ChemDraw or any drawing package that can produce a Molfile. When the user selects a search form object that is associated with alphanumeric data, PCC displays a dialog that allows the user to define query terms using operators that can be picked from a dropdown list, thus eliminating problems with different query syntax for text searches in different data sources.

For structure searches PCC uses the Chemical Development Kit (CDK) [1]. For the out-of-the box SDFile loader we use Microsoft Excel as Molfile and data container. If no Excel is available the user can use a comma separated file as option.

In addition to saving search forms, PCC allows to import/export search results to list files and to generate HTML files from search results, including structures, for printing. Structures are encoded in the HTML file using the free ChemView ActiveX control from Hyleos.net [2].

PCC is unique in that the structure information can reside in comma separated files (CSV), Microsoft Excel, Oracle, Microsoft SQLServer, MySQL and Microsoft Access and the data can reside in any of the aforementioned systems. This gives maximum integration possibilities to add structures to existing data sources. We recommend using our consulting expertise to adjust the system to existing data sources.

You can download PCC at <http://www.akosgmbh.eu/pcc>.

References

1. . Chemical Development Kit (CDK), is a library of Java classes for chemo-, bioinformatics, computational chemistry, and chemometrics. The CDK is available for download from <http://sourceforge.net/projects/cdk>
2. . ChemViewX ActiveX control is a free MDL Molfile renderer from Hyleos.net. ChemViewX is available for download from <http://www.hyleos.net>.