

Oral presentation

Open Access

Breaking the language barrier: chemical nomenclature around the globe

R Sayle

Address: OpenEye Scientific Software, Suite #D, 9 Bisbee Court, Santa Fe, New Mexico, NM 87508, USA
from 3rd German Conference on Chemoinformatics
Goslar, Germany. 11-13 November 2007

Published: 26 March 2008

Chemistry Central Journal 2008, 2(Suppl 1):S11 doi:10.1186/1752-153X-2-S1-S11

This abstract is available from: <http://www.journal.chemistrycentral.com/content/2/S1/S11>

© 2008 Sayle

The use of chemical compound names remains the primary method for conveying molecular structures between chemists and researchers. In research articles, patents, chemical catalogues, government legislation and textbooks, the use of IUPAC and traditional compound names is universal, despite efforts to introduce more machine-friendly representations such as identifiers and line notations. Fortunately, advances in computing power now allows chemical names to be parsed and generated (read and written) with almost the same ease as conventional connection tables. A significant complication, however, is that although the vast majority of chemistry uses English nomenclature, a significant fraction is in other languages. This complicates the task of filing and analyzing chemical patents, purchasing from compound vendors and text mining research articles or web pages. This talk describes the issues and complications with manipulating compound names in various languages, such as Japanese, and describes current state-of-the-art tools to simplify the process, including recent work at OpenEye to support reading and writing of German chemical nomenclature.

References

1. "Rechtsvorschriften für Gefährliche Stoffe: Einstufung und Kennzeichnung in der Europäischen Gemeinschaft". 1987.
2. Hellwinkel D: "Die systematische Nomenklatur der organischen Chemie: Eine Gebrauchsanweisung". Springer, 2005.