

ORAL PRESENTATION

Open Access

Semantics vs. statistics in chemical markup

Colin Batchelor

From 7th German Conference on Chemoinformatics: 25 CIC-Workshop
Goslar, Germany. 6-8 November 2011

Since the late 1990s, natural language processing (NLP) has seen a massive shift from high-precision, low-recall systems based on small sets of hand-written rules, to methods based on the statistical analysis of large corpora. The field of chemoinformatics, likewise, is dominated by statistical and machine-learning approaches. In recent years, however, pharmaceutical companies have been engaging more and more with Semantic Web technologies, which are largely built around the sorts of hand-written systems that NLP has moved away from this century. We discuss where our current text analysis and Semantic Web efforts at the Royal Society of Chemistry are headed and how we're making use of the unreasonable effectiveness of data.

Published: 1 May 2012

doi:10.1186/1758-2946-4-S1-O16

Cite this article as: Batchelor: Semantics vs. statistics in chemical markup. *Journal of Cheminformatics* 2012 **4**(Suppl 1):O16.

Publish with **ChemistryCentral** and every scientist can read your work free of charge

"Open access provides opportunities to our colleagues in other parts of the globe, by allowing anyone to view the content free of charge."

W. Jeffery Hurst, The Hershey Company.

- available free of charge to the entire scientific community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
<http://www.chemistrycentral.com/manuscript/>



ChemistryCentral

Correspondence: batchelor@rsc.org
Royal Society of Chemistry, Thomas Graham House, Cambridge, CB4 0WF,
UK



© 2012 Batchelor; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.