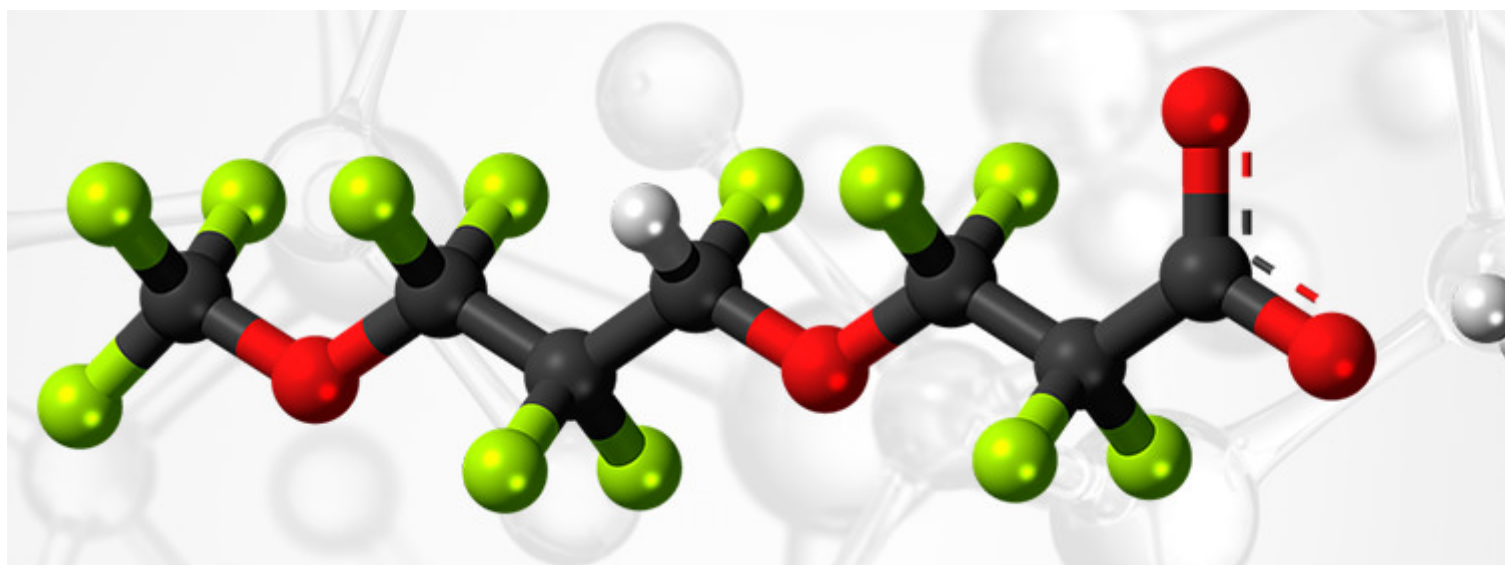


[Inicio](#) > Alfonso Valencia reviews the state-of-the-art of Information Retrieval and Text Mining Technologies for Chemistry

---

## [Alfonso Valencia reviews the state-of-the-art of Information Retrieval and Text Mining Technologies for Chemistry](#)

Efficient access to chemical information contained in scientific literature, patents, technical reports, or the web is a pressing need shared by researchers and patent attorneys from different disciplines.



A Review authored by the BSC Life Sciences Department, Alfonso Valencia, provide a comprehensive and in-depth description of the stat-of-the-art of Information Retrieval and Text Mining Technologies for Chemistry.

The review has been published by [American Chemical Society](#) Publications:  
[pubs.acs.org/doi/abs/10.1021/acs.chemrev.6b00851](https://pubs.acs.org/doi/abs/10.1021/acs.chemrev.6b00851)

### **Abstract**

Efficient access to chemical information contained in scientific literature, patents, technical reports, or the web is a pressing need shared by researchers and patent attorneys from different chemical disciplines. Retrieval of important chemical information in most cases starts with finding relevant documents for a particular chemical compound or family. Targeted retrieval of chemical documents is closely connected to the automatic recognition of chemical entities in the text, which commonly involves the extraction of the entire list of chemicals mentioned in a document, including any associated information. In this Review, we provide a comprehensive and in-depth description of fundamental concepts, technical implementations, and current technologies for meeting these information demands. A strong focus is placed on community challenges addressing systems performance, more particularly CHEMDNER and CHEMDNER patents tasks of BioCreative IV and V, respectively.

Considering the growing interest in the construction of automatically annotated chemical knowledge bases that integrate chemical information and biological data, cheminformatics approaches for mapping the extracted chemical names into chemical structures and their subsequent annotation together with text mining applications for linking chemistry with biological information are also presented. Finally, future trends and current challenges are highlighted as a roadmap proposal for research in this emerging field.

**DOI:** 10.1021/acs.chemrev.6b00851

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---

**Source URL (retrieved on 15 Jul 2024 - 18:47):** <https://www.bsc.es/es/news/bsc-news/alfonso-valencia-reviews-the-state-the-art-information-retrieval-and-text-mining-technologies>