

The following review is the first in a series of articles surveying free online resources of potential interest to chemists. The author plans to cover general resources, chemical informatics, mathematical applications, and journals and publications.

Free Information Resources for Chemists

by Leslie Glasser

As chemists, we need generic software (e.g., Web browsers, word processors, spreadsheets, and general graphics programs) as well as specialized information and software (e.g., for naming compounds, communicating and receiving chemical information in the form of structures and data, accessing databases, and performing symbolic mathematics). Although there is considerable commercial software available to meet all these requirements, this article focuses on specialized chemical software that is currently available free for the taking on the Internet. Because the scope of this topic is vast and is changing continuously, only some of the more well-known or potentially valuable products will be mentioned. Although much of the material discussed in this article is free across the board, some is free only for noncommercial use.

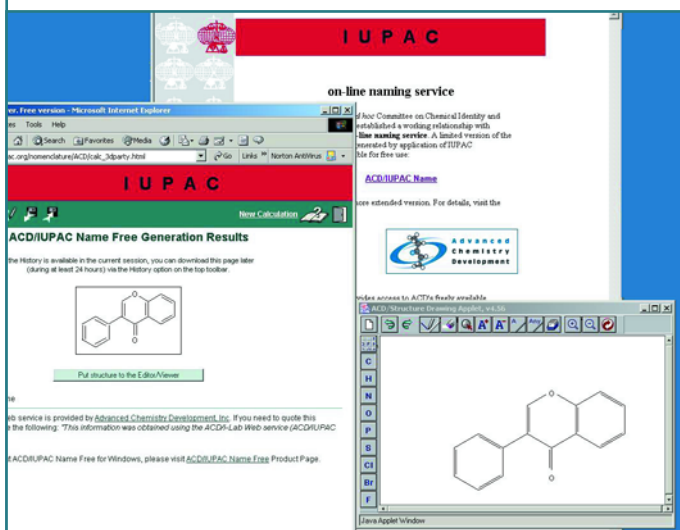
A comprehensive and useful general source of information is the textbook *Chemoinformatics*¹ and its related *Handbook*.² The website www2.chemie.uni-erlangen.de/publications/ci-book/tb_websites.html provides links to a number of sources (although the site was last updated 19 October 2003). The U.S. National Institute of Standards & Technology (NIST) provides access to many free database resources www.nist.gov/srd and lists internationally recommended values of fundamental physical constants at physics.nist.gov/cuu/Constants. ChemIndustry.com labels itself as "the worldwide search engine of the chemical industry" and provides an extensive chemical listing and searching facility especially (but not solely) directed toward industrial application. One of their recent newsletters www.chemindustry.com/newsletter lists "Chemistry Software Resources" (Number 57) and "Free Chemical Information Resources, Parts 1, 2, and 3" (Numbers 58-60). A further important link to a wide range of commercial scientific material is www.fiz-informationsdienste.de.

Generic Products

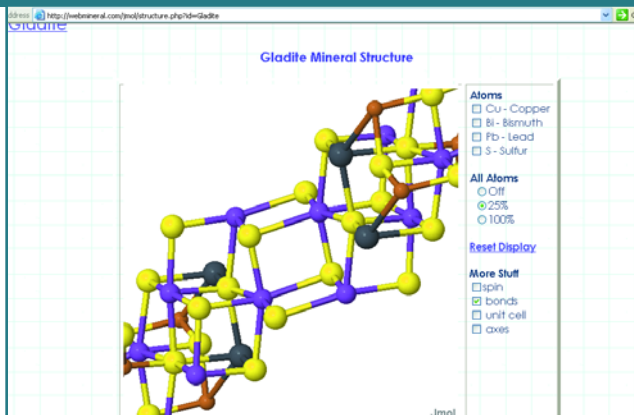
The Sheffield ChemDex www.chemdex.org functions as a directory to chemistry on the Web (e.g., websites, chemistry departments, and chemical information), while the Royal Society of Chemistry www.chemsoc.org provides a general chemistry resource.

ACD/IUPAC Name Free 8.05 is a free application (with some imposed limitations) available from the IUPAC website that enables users to name chemical structures according to IUPAC Recommendations www.iupac.org/nomenclature/ACD/calc_3dparty.html. This service is provided jointly by IUPAC and Advanced Chemistry Development, Inc. The full commercial version is available from ACD/Labs itself.

An important recent development is the IUPAC International Chemical Identifier (InChI™) www.iupac.org/inchi, which is a text-based, non-proprietary, unique identifier for chemical substances that can be used in printed and electronic data sources, thus enabling easier linking of diverse data compilations. InChI names have already been associated with tens of millions of chemical entities, thus rendering searching the Web for references to such materials as simple as any other text search.³ InChI can be generated online on the website pubchem.ncbi.nlm.nih.gov/edit (scroll down to



Access the ACD/IUPAC online naming service at www.iupac.org/nomenclature.



The mineral structure of gladite as seen on webmineral.com.

choose from SMILES, SMARTS, or InChI). Alternatively, and for local use, a copy of the InChI software can be downloaded from the IUPAC website.

Links to published scientific material (which itself may not always be free) can be found using Google Scholar <scholar.google.com> and PubMed <www.pubmed.gov>, which provides links to about 16 million citations. Other literature search engines include Scirus <www.scirus.com>, Scopus <www.scopus.com>, and ScienceDirect <www.sciencedirect.com> (ScienceDirect has limitations for guest users). Chmoogle <www.chmoogle.com> provides many links to sources of materials. The Google Science Directory <www.google.com/Top/Science/Chemistry> provides links to nearly 5 500 chemistry-related sites—useful for broad assessments.

An important and reliable source of chemical data is NIST's Chemistry Webbook <webbook.nist.gov/chemistry>. WebElements <www.webelements.com> provides detailed information on the chemical elements. A comprehensive listing of mineral information is at <webmineral.com>. Structural mineral databases are maintained at <database.iem.ac.ru/mincryst> and <www.minsocam.org/MSA/Crystal_Database.html>. Protein structural information can be obtained from the Protein Databank <www.rcsb.org/pdb>.

References

1. *Chemoinformatics: A Textbook*, Gasteiger, J; Engel, T., eds., Wiley-VCH, Weinheim, Germany, 2003.
2. *Handbook of Chemoinformatics: From Data to Knowledge*, Gasteiger, J., ed., Wiley-VCH, Weinheim, Germany, 2003, 4 vols.
3. The utility of the selectivity of an InChI search in Google is shown by the fact that a search for "ethanol" yields 12.4 million hits, whereas a search for "InChI=1/C2H6O/c1-2-3/h3H,2H2,1H3" yields fewer than 100 hits, including an elaborate description and discussion on Wikipedia

<en.wikipedia.org/wiki/Wikipedia:About>, the collaborative—hence nonauthoritative—online encyclopedia.

Leslie Glasser <leslieglasser@yahoo.co.uk> is chairman of the IUPAC Committee on Printed and Electronic Publication (CPEP). He is a professor in the Department of Applied Chemistry, Nanochemistry Research Institute, of the Curtin University of Technology, in Perth, Australia.

TOXNET—Information Resources and Services in Toxicology

by John Duffus

TOXNET (TOXicology data NETwork), a Web portal established by the United States National Library of Medicine (NLM) <<http://toxnet.nlm.nih.gov>>, is a cluster of databases covering toxicology, hazardous chemicals, environmental health, and related areas. It is managed by the Toxicology and Environmental Health Information Program in the Division of Specialized Information Services of NLM.

TOXNET provides free access to and easy searching of the following databases:

HSDB® (Hazardous Substances Data Bank)

A factual database focusing on the toxicology of over 4 900 potentially hazardous chemicals.

IRIS (Integrated Risk Information System)

A database from the U.S. Environmental Protection Agency (EPA) containing carcinogenic and noncarcinogenic health risk information on over 500 chemicals.

ITER (International Toxicity Estimates for Risk)

This database contains data in support of human health risk assessments. It is compiled by Toxicology Excellence for Risk Assessment, and contains over 600 chemical records.

CCRIS (Chemical Carcinogenesis Research Information System)

A scientifically evaluated and fully referenced data bank, developed and maintained by the National Cancer Institute. It contains over 8 900 chemical records with carcinogenicity, mutagenicity, tumor promotion, and tumor inhibition test results.