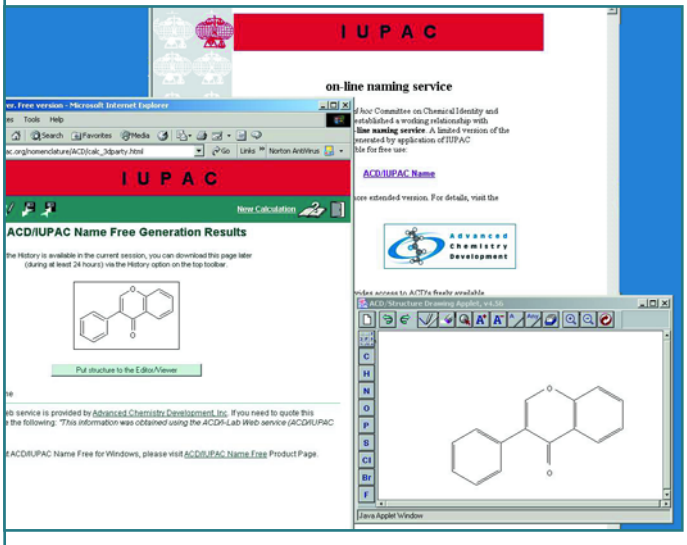
The following review is the first in a series of articles surveying free online resources of potential interest to chemists. The author plans to cover general resources, chemical informatics, mathematical applications, and journals and publications.

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.



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

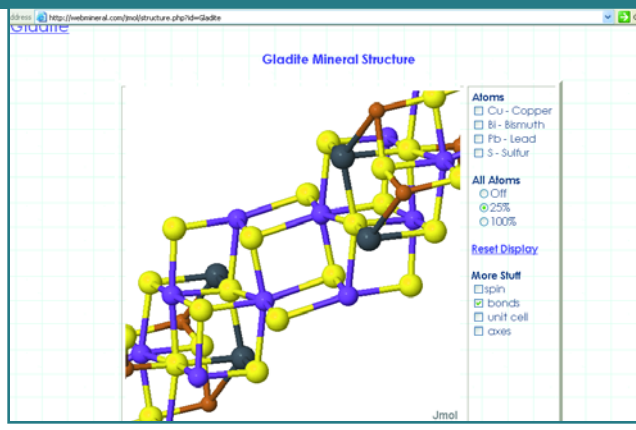
A comprehensive and useful general source of information is the textbook *Chemoinformatics*[1] and its related *Handbook*.[2] The website <www2.chemie.uni-erlangen.de/publications/ci-book/tb_websites.html> provides links to a number of sources (although the site was last updated 19 October 2003). The 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, nonproprietary, 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



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 "InChI=1/C2H6O/c1-2-3/h3H,2H2,1H3" 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 <lesliеglasser@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.

Internet Connection

GENE-TOX (Genetic Toxicology)
A toxicology database created by the U.S. EPA containing genetic toxicology test results on over 3 000 chemicals.

Tox Town
An interactive guide to commonly encountered toxic substances, your health, and the environment.

Household Products Database
This database provides information on the potential health effects of chemicals contained in more than 5 000 common household products used inside and around the home.

Haz-Map®
An occupational toxicology database designed primarily for health and safety professionals, but also for consumers seeking information about the health effects of exposure to chemicals and biologicals at work.

Drugs and Lactation Database (LactMed)
A peer-reviewed and fully referenced database of drugs to which breastfeeding mothers may be exposed. Among the data included are maternal and infant levels of drugs, possible effects on breastfed infants and on lactation, and alternate drugs to consider.

TOXMAP
A website from the National Library of Medicine (NLM) that uses maps of the United States to show the amount and location of toxic chemicals released into the environment.

TOXLINE®
A bibliographic database providing comprehensive coverage of the biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals from 1965 to the present.

DART®/ETIC (Development and Reproductive Toxicology/Environmental Teratology Information Center)
A bibliographic database covering literature on reproductive and developmental toxicology. DART is managed by NLM and funded by the EPA, the National Institute of Environmental Health Sciences and NLM. DART/ETIC contains references to reproductive and developmental toxicology literature published since 1965.

Toxics Release Inventory
A series of databases that describe the releases of toxic chemicals into the environment annually for the 1987–2004 reporting years.

ChemIDplus

A database providing access to structure and nomenclature authority databases used for the identification of chemical substances cited in NLM databases.

TOXNET databases are accessible free of charge at <http://toxnet.nlm.nih.gov>. Links are available from TOXNET to PubMed®, NLM's free Web interface to the world's biomedical literature, and to additional sources of toxicological information. These may be searched through ToxSeek, <http://toxSeek.nlm.nih.gov>, which is a meta-search engine that enables simultaneous searching of many different information resources on the Internet. Associated with these websites is WebWISER <http://wiser.nlm.nih.gov>, a system designed to assist first responders in hazardous material incidents. WISER provides a wide range of information on hazardous substances, including substance identification support, physical characteristics, human health information, and containment and suppression advice.

Since its inception, TOXNET has offered users a tutorial website, ToxTutor, in order to help them make the most of the available data. Recently, it became clear that this had to be updated and a project, ToxLearn, was started in order to update and expand the content of ToxTutor using current programming technology. The ToxLearn project is sponsored by the NLM and, as a U.S. government agency, all materials they develop are in the public domain. There are no ownership rights. NLM is providing support for the project, including staff time (Dr. Philip Wexler), and is underwriting the technological Web interface for the content through the Web developer, Patient Education Institute. The final product will be resident on the NLM server and maintained there. The *IUPAC Glossary of Terms Used in Toxicology* is integrated into the ToxLearn package and hyperlinked to text where appropriate. On this NLM server, the glossary is accessible at <http://sis.nlm.nih.gov/enviro/glossary-main.html> (see box).

The ToxLearn project plan has two phases. Phase 1 was the production of text and design of a pilot of the technology interface for the rest of the project. Dr. Mike Kamrin, Michigan State University (modules 1 and 2), and Dr. Paul Wright, RMIT University (module 3), are the authors of the pilot modules. ToxLearn has been adopted as a project of the U.S. Society of Toxicology (SOT) Education Committee, which is funding the project. SOT will soon issue an open call to authors for the phase 2 modules. The ToxLearn Working Group, which includes the author of this article, will recommend suitable authors chosen from those who apply. The Education Committee will make the final decision on their appointment.

John Duffus <j.h.duffus@blueyonder.co.uk> is the chairman of the Subcommittee on Toxicology and Risk Assessment and a titular member of the Chemistry and Human Health Division of IUPAC.

http://toxnet.nlm.nih.gov

IUPAC Glossary of Terms Used in Toxicology

The *IUPAC Glossary of Terms Used in Toxicology* was first published in *Pure and Applied Chemistry* in 1993[1] and has been supplemented since by the *Glossary of Terms Used in Toxicokinetics*[2] and the *Explanatory Dictionary of Concepts in Toxicology*. A revised and expanded version is currently under public review (see p. 23), and more than 260 terms have been added to it. The original 1993 glossary was the basis for the glossary incorporated in the textbook *Fundamental Toxicology for Chemists*, prepared with IUPAC support and published by the Royal Society of Chemistry.[3] A second edition of this textbook simply titled *Fundamental Toxicology*[4] was recently published by the Royal Society of Chemistry and includes an extended version of the original glossary (see Bookworm, p. 24). Another extended version has also been included in the *Encyclopedia of Toxicology*.[5]

References

1. J. H. Duffus, P*ure Appl. Chem.* **65**, 2003–2122 (1993).
2. M. Nordberg, J. H. Duffus, and D. M. Templeton, *Pure Appl. Chem.* **76**, 1033–1082 (2004).
3. J. H. Duffus and H. G. J. Worth (eds.), *Fundamental Toxicology for Chemists*, Pub. Royal Society of Chemistry, Cambridge (1996).
4. J. H. Duffus and H. G. J. Worth (eds.), *Fundamental Toxicology*, 2nd Edition, Pub. Royal Society of Chemistry, Cambridge (2006).
5. P. Wexler (ed.), *Encyclopedia of Toxicology*, 2nd Edition, Pub. Academic Press, San Diego (2005).