

The [DIGITAL]  
Chronicle:  
\$1.22  
an issue

THE CHRONICLE OF HIGHER EDUCATION

## Research & Publishing

From the issue dated June 29, 2001

### Chemists See More Data, but Not the Deluge Experienced by Other Scientists

By LILA GUTERMAN

Chemists call their field "the central science," but in one way, at least, they're sidelined. While scientists

in many other fields share data through freely available online repositories, chemists seem uninterested. As a result, they haven't been hit by a deluge of information, as other scientists have, but instead have experienced a steady increase.

#### ALSO SEE:

[Learning to Swim in the Rising Tide of Scientific Data](#)

"Chemistry has always been different," says Stephen R. Heller, a retired government chemist. "Not necessarily better or worse, just different."

That difference shows up in the field's most venerable database, the Chemical Abstracts Service, which contains summaries of 18 million journal papers. Set up almost 100 years ago, the abstracts service is a profitable enterprise -- a fact that may have set the tone for all large chemical databases.

For online access to Chemical Abstracts Service, universities that offer Ph.D.'s pay the American Chemical Society at least \$27,800 a year; undergraduate institutions pay at least \$13,900. In return, their students and scientists can search databases of compounds, reactions, and literature citations.

Another database that many chemists cite as important to their work, Beilstein, is sold by MDL Information Systems, a company that specializes in chemical information and software to manage it. Online access to chemical compounds, reactions, and associated scientific data costs around \$5,000 for undergraduate colleges and from \$7,000 to \$27,500 for Ph.D.-granting universities, depending on their size, according to a company official, who later said the prices vary and would not give further details.

When scientists are accustomed to paying for information, says Mr. Heller, "it becomes a little more difficult to give it away -- or to get

Search The Site

[More options](#) | [Back issues](#)

[Home](#)

[News](#)

[Today's news](#) [Current](#)

[issue](#)

[Special issues & data](#) ⊕

[The Faculty](#) [Research &](#)

[Books](#) [Government &](#)

[Politics](#) [Money &](#)

[Management](#)

[Information](#)

[Technology](#) [Students](#)

[Athletics](#) [International](#)

[Community Colleges](#)

[Short Subjects](#) [Gazette](#)

[Corrections](#)

[Opinion & Forums](#) ▶

[Careers](#) ▶

[Multimedia](#)

[Leadership Forum](#)

[Technology Forum](#)

[Resource Center](#)

[Campus Viewpoints](#)

[Services](#)

[Help](#) [Contact us](#) [Subscribe](#)

[Day pass](#) [Manage your](#)

[account](#) [Advertise with us](#)

[Rights & permissions](#)

[Employment Opportunities](#)

/r



CruiserAlert™

See a Demo.

Request a  
White Paper.

TRY IT OUT ▶



HIGH-PRIORITY &  
EMERGENCY  
COMMUNICATION

support for giving it away."

Complicating the situation further is chemistry's bonds to industry. For decades, the discipline has had an emphasis on creating marketable products, such as pharmaceuticals or plastics. "People buy chemicals all the time," says Gary Mallard, the group leader of chemical reference data at the National Institute of Standards and Technology. In contrast, less of a market exists for work by physicists or other scientists, he says. "Who in the world buys fundamental physical constants?"

Mr. Heller believes that the trend makes chemists leery of sharing their data. "If someone comes up with something, it could very well have considerable intellectual-property value," he says.

Still, some chemists face the same potential to drown in data as do biologists handling information about the genome or earth scientists analyzing data pouring down from a satellite. A recently developed method, called combinatorial chemistry, allows chemists to make thousands of compounds at once and then test them for various properties.

Yet the fear of data overload is more limited in chemistry than in other fields because the technique is most often used in industry. So the resulting database is "all internal to a company. The compounds are their intellectual property," says David A. Evans, a product manager for MDL Information Systems.

Of course, some exceptions exist. Chemists studying biological molecules use many of the same databases that biologists do, such as GenBank. And Mr. Mallard works on NIST's Chemistry WebBook, an online repository that contains physical and chemical data, including spectra of various sorts, on some 36,000 chemical compounds. The site gets hits from only around 110 distinct users a day, a far cry from GenBank's 20,000 daily users. Chemistry WebBook's small size may limit its usefulness to scientists, Mr. Heller says: "It pales in comparison to the 10 or 15 million [compounds] in Beilstein or Chemical Abstracts."

Chemists have shared data about the structure of crystals through a nonprofit data center at the University of Cambridge since 1965. That databank holds information about approximately 210,000 compounds. Unlike the genome or satellite databases, which experience huge and constant inflows of information, the Cambridge database is "not expecting someone to dump a million structures a day," says Jeremy Frey, a senior lecturer in chemistry at the University of Southampton.

He thinks one reason chemistry has not established large, freely available data sets is that "big science" has never gained a foothold in the discipline. "We don't have huge projects with only one objective," he says. "We're all busily doing different things."

---

<http://chronicle.com.libezproxy2.syr.edu>  
Section: Research & Publishing  
Page: A15

---



[Easy-to-print](#) version



[E-mail](#) this article

---

[Copyright](#) © 2001 by The Chronicle of Higher Education