June 20, 2005

Dear Colleague:

I want to inform you about a very important issue that could have a significant impact on ACS and science.

It centers on a database launched by NIH last October called PubChem. You may have read about this subject in the June 13 issue of *Chemical & Engineering News* [http://pubs.acs.org/cen/government/83/8324gov1.html](http://pubs.acs.org/cen/government/83/8324gov1.html) and other places. Allow me to give you some additional context and also to address some commentary that has appeared in various reports.

One of the principal reasons NIH launched PubChem is to disseminate data that will result from its Molecular Libraries Initiative ([http://nihroadmap.nih.gov](http://nihroadmap.nih.gov)). ACS has supported this effort from the outset and has offered to work with NIH in this regard. Contrary to some reports, ACS is not calling for a termination of PubChem. We have asked, however, that PubChem not significantly duplicate the CAS Registry® – the world's hallmark database of 25 million chemical substances. This is the cause of our concern about the current situation.

On June 16, the ACS position on PubChem was reinforced when the House of Representatives Appropriations Committee expressed similar concern about PubChem replicating scientific information services. The committee urged NIH to work with private sector providers to properly focus PubChem to avoid unnecessary duplication and competition with private sector chemical databases.

I am pleased to say that both ACS and NIH want a “win-win” situation between two organizations whose mission is the advancement of science. ACS is working to come to a resolution soon and I am optimistic.

Having said that, I want to explain in more detail why ACS is concerned about the current situation.

At this time, the PubChem database contains no information from the Molecular Libraries Initiative. NIH has only just awarded the contracts for that initiative and it will be some time before molecules from that initiative appear in PubChem. In the meantime, the PubChem database contains somewhat less than a million substances. NIH has stated that it will grow this database substantially as it seeks to make it as comprehensive as possible. In addition to re-collecting substances already available through the CAS Registry, PubChem disseminates them through a platform of information fields and search technology that mimics the Registry. ACS is very concerned that as NIH populates PubChem with millions more substances though this similar platform, it will be providing, for “free,” a taxpayer-subsidized substitute for...
the CAS Registry. And while ACS continues to work with NIH toward a reasonable parameter to avoid unnecessary duplication, NIH maintains that it must be allowed to expand PubChem in any direction.

By way of background, the CAS Registry is renowned for its accuracy and comprehensiveness. It is used as a stand-alone search tool by thousands of scientists around the world. Nearly 40% of all revenues from CAS are directly attributable to searches in the CAS Registry alone (without linking to other CAS databases). In addition, the CAS Registry is the underpinning for many of the state-of-the-art desktop scientific information tools that CAS has developed over decades, such as SciFinder® and SciFinder Scholar™. Together, these tools have compressed what would formerly take weeks or months of research time into minutes or seconds—literally fast-forwarding scientific progress.

In developing the CAS Registry, ACS benefited from small seed grants from the National Science Foundation, which helped develop the underlying technology. Since then, ACS has invested more than $500 million to develop, maintain, and enhance the CAS Registry. ACS continues to use its own financial resources and the skills of hundreds of highly skilled CAS scientists to keep the CAS Registry up-to-date. The costs are borne by the users, which is a long-standing and effective operational model.

When the government builds a major new database to replicate scientific data and provide it for free to researchers who already have it, we must carefully consider the full short- and long-term financial costs and consequences to science:

- If CAS and other private services are displaced, the costs and responsibilities of the Federal government will grow. And the quality, management, and long-term preservation of scientific information could suffer as it is subjected to changing agency priorities and reduced annual budgets in this era of budget deficits. ACS questions the premise that the Federal government should be the funder, publisher, and repository of all scientific information regardless of whether it was federally funded. ACS has shown a long-term commitment to the stewardship of the currently robust scientific literature.

- A rapidly growing, tax-subsidized PubChem with no boundaries constitutes unfair government competition and runs counter to long-standing Office of Management and Budget circulars that direct agencies to leverage and rely on, not re-create and compete improperly with, similar private services.

- Duplication is inherently inefficient, and the full, out-year costs of PubChem will be substantial. In a time of flat budgets, when we would all like to see more basic research funding, we do not believe that duplicating a highly respected database is a good—or proper—use of government resources.

The issue is not whether to advance the NIH Molecular Libraries Initiative but how. ACS has been advancing science for more than 100 years and would never suggest a policy that would adversely affect scientific progress. The ACS believes that a cooperative agreement between NIH and ACS can accelerate the progress of the MLI while insuring the integrity of the CAS Registry for the future.

In fact, ACS has a long history of cooperation with NIH that includes mobilizing ACS presidents, the ACS Board of Directors, and ACS members to advocate for a doubling of the NIH budget. Together our organizations can continue to do great things to advance chemistry, biology, and scientific discovery.

The increase and diffusion of chemical knowledge is the cornerstone of the ACS mission and its Congressional Charter. ACS has shown its unwavering commitment to this mission for nearly 130 years,
through its high-quality journals and CAS databases and research tools. CAS products such as SciFinder and SciFinder Scholar are mainstays in universities, corporations, and government labs around the globe. More than half of the Society’s net revenues are generated by CAS. All but a fraction of one percent is reinvested back into supporting the costs of ACS publishing programs and innovation in these programs. Any net revenues are also used to provide funding for a wide array of ACS services to ACS members and the profession as a whole.

As an example of this innovation, CAS is introducing a new data mining tool next month that will help scientists glean even more information from the ever-increasing reams of information. Later this year, CAS will introduce CAS Mobile™, which allows users to conduct complex searches from PDA’s—a first for scientific information retrieval. New and important features in SciFinder are also being introduced later this year. This is the kind of innovation we have come to depend upon and we cannot afford to take for granted. A serious reduction in revenues from CAS will have immediate consequences for the viability of CAS database publishing operations and also for ACS’s ability to continue many of our programs aimed at fulfilling our charter to advance the science and profession.

Last, let me put a human face on this issue. In Columbus, Ohio, CAS employs 1,300 highly skilled chemists and other scientists, technologists, and support staff. These are the talented men and women who have devoted their lives to ensure that the data in all of the CAS databases are comprehensive, accurate, and timely. They not only build the CAS Registry from all publicly disclosed literature and patents, they also expertly index those documents for retrieval and analysis, and develop the chemistry-focused research tools that take best advantage of these world-leading databases. They ensure that a state-of-the-art data center makes SciFinder, SciFinder Scholar and STN available to the world’s scientific community around the clock. At a time when the U.S. is struggling to retain high-technology jobs in the U.S. and Congressmen are worrying about the competitiveness of the U.S. in science and technology, it seems counter-productive for the Federal government to be engaged in policies that could threaten 1,300 jobs in Ohio.

I hope I have made the ACS position clear and want to emphasize again that ACS is committed to working with NIH to resolve this issue in a speedy and amicable fashion that benefits science and scientists. If you have additional questions, please contact me at president@acs.org.

Sincerely yours,

Dr. William F. Carroll, Jr.
President, American Chemical Society