

# New Web-based system leads to better, more timely data

October 27 2006

---

After two years of work, an innovative project using Web-based technologies to speed researcher access to a large body of new scientific data has demonstrated that not only access to but also the quality of the data has improved markedly. A new paper on the Web-enabled ThermoML thermodynamics global data exchange standard notes that the data-entry process catches and corrects data errors in roughly 10 percent of journal articles entered in the system.

A landmark partnership between the National Institute of Standards and Technology (NIST), several major scientific journals and the International Union of Pure and Applied Chemistry (IUPAC), ThermoML was developed to deal with the explosive growth in published data on thermodynamics.

Thermodynamics is essential to understanding and designing chemical reactions in everything from huge industrial chemical plants to the biochemistry of individual cells in the body. With improvements in measurement technology, the quantity of published thermophysical and thermochemical data has been almost doubling every 10 years.

This vast flood of information not only presents a basic problem for researchers and engineers--how to find the data they need when they need it.-- but also has strained the traditional scientific peer-review and validation process. "Despite the peer-review process, problems in data validation have led, in many instances, to publication of data that are grossly erroneous and, at times, inconsistent with the fundamental laws

of nature," the authors note.

The ThermoML project began as an attempt to simplify and speed the delivery of new thermodynamic data from producers to users. The system has three major components -- ThermoML itself, an IUPAC data format standard based on XML (a generic data formatting standard) customized for storing thermodynamic data; Software tools developed at the NIST Thermodynamic Research Center (TRC) to simplify entering data into the system in formats close to those used by the original journal documents, displaying it in various formats and performing basic data integrity checks; and The ThermoData Engine, a sophisticated expert system developed at NIST, that can generate on demand recommended, evaluated data based on the existing experimental and predicted data and their uncertainties.

Authors writing for five major journals that are partners in the program, the Journal of Chemical and Engineering Data, the Journal of Chemical Thermodynamics, Fluid Phase Equilibria, Thermochemica Acta, and the International Journal of Thermophysics, participate in the process by submitting the data for their articles using GDC software (available from NIST). The data are evaluated, and any potential inconsistencies reported back to the authors for verification. Based on two years of experience and some 1,000 articles, the authors write, an estimated 10 percent of articles reporting experimental thermodynamic data for organic compounds contain some erroneous information that would be "extremely difficult" to detect through the normal peer-review process.

Citation: M. Frenkel et al. New global communication process in thermodynamics: impact on quality of published experimental data. J. Chem. Inf. Model. ASAP Article. Web Release Date: October 11, 2006. [pubs.acs.org/cgi-bin/abstract....1/abs/je025645o.html](http://pubs.acs.org/cgi-bin/abstract....1/abs/je025645o.html)

Source: NIST

Citation: New Web-based system leads to better, more timely data (2006, October 27) retrieved 25 April 2024 from <https://phys.org/news/2006-10-web-based.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.