

ACS/CODATA Symposium on Numeric/Factual Materials and Chemical Databases

J. G. Kaufman

MPD Network Inc., Chemical Abstracts Service, 2540 Olentangy River Road, Columbus, Ohio 43210

Received November 3, 1992

Providing chemists and other scientists and engineers with direct access to quantitative performance and properties data has taken on increased importance in recent years. At the April 1992 ACS Spring Meeting, the Chemical Information Division of the ACS focused a full-day symposium on that subject. Sponsorship of the sessions was shared by CODATA, and Dr. Gordon Wood, Secretary-General of CODATA, and I cochaired the meeting.

The 10 papers that follow constitute the proceedings of those sessions and provide a good overview of both the scope and importance of the activity and of the progress that has been made in the past 10 years since the Fairfield Glade Conference¹ brought a sharp focus to the need. CODATA was a sponsor of that original meeting, and the American Chemical Society has joined with the National Material Property Data Network, Inc., in providing an important part of the response to the need. Thus, it is fitting that these two organizations cooperated in bringing this conference and the following "proceedings" to light.

In the initial paper in the series, I set the stage for what follows by describing the critical features of numeric data, its focus and complexities, in terms of users needs, and illustrate why such resources are expensive to provide. The complexity of numeric data is further highlighted in Westbrook's paper on the technology and difficulties of data capture from tables and figures when converting hardcopy handbooks and publications to electronic databases.

Gragg follows with an accounting of the role of standards in this activity; the leadership role of ASTM Committee E49 on Computerization of Materials and Chemical Property Data is described, both the accomplishments and the remaining needs. Barrett, at the time Chairman of CODATA's Task Group on Materials Database Management, provides a sense of the socioeconomic implications of numeric data and, in a classic example, illustrates the importance of quality and reliability in numeric data.

One of the key features of identification of substances in databases both for storage and for search and retrieval is the

Chemical Abstracts Service (CAS) Registry Number (RN), and Moulton's paper catalogues a number of enhancements made to that numbering system in recent years to increase its precision and usefulness for various material systems.

Wood and Ho describe the responsibilities and technical approach to managing important data centers: the Canadian Institute of Scientific and Technical Information (CISTI) and the Center for Information and Data Analysis and Synthesis (CINDAS), respectively. The importance and methodologies of data evaluation are illustrated. Another approach to providing direct access to materials data for scientists and engineers is described in the Drago/Kaufman paper about online networking of materials and chemical databases.

The diversity of types and media for numeric/factual data are illustrated by Barth's paper on an online database covering spectral and chemical structure data and by Staley's paper on a personal computer-based approach to X-ray diffraction data.

For other important work in this area, we refer you to the proceedings of the ASTM Symposia on Computerization and Networking of Material Property Data² and to CODATA Conference proceedings (e.g., The Role of Data in Scientific Progress.³) and various issues of the CODATA Bulletin (notably No. 69).

REFERENCES AND NOTES

- (1) *Computerized Materials Data Systems: The Proceedings of a Workshop at Fairfield Glade, TN, Nov 7-11, 1982*; Office of Standard Reference Data, National Bureau of Standards: Gaithersburg, MD, 1983.
- (2) *ASTM Spec. Tech. Publ. Nos. 1017 and 1106*.
- (3) Published by North-Holland Publishing Co., Amsterdam.