INTERNATIONAL COUNCIL OF SCIENTIFIC UNIONS
ISSN 0538-6918

COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY

63 CODATA / NEWSLETTER

MAY 1993

CODATA '94 — Data and Knowledge in a Changing World

The Quest for a Healthier Environment

INVITATION

The International Conference 18-22 September 1994, in Chambéry, France, will feature scientific and technical data challenges in the traditional sectors of industry, physics and chemistry as well as in bio-, geo- and space sciences. The Conference will be centered around three main symposia:

— Materials

— Environment and

— Communication

These symposia will identify interdisciplinary problems and synthesize existing information to highlight crucial research needs and solutions.

Materials and Thermodynamics will deal with classical and recent problems of materials and matter in various fields, such as:

- classical and novel materials including compound modelling
- computer-aided characterization and quality control
- knowledge bases and materials databases
- materials data for innovation, composite materials design, and industrial decision making

This symposium, organized in conjunction with the *Centre Européen de Réflection et d'Etude en Thermodynamique (CERET)*, will stress the importance of materials and matter in different states: pure compounds, solutions, mixtures (equilibrium and non-equilibrium patterns), new products, stabilized- and stable structures, thermotransient data, etc.

Environment and Toxicity will seek to elucidate environmental problems on a general and regional level considering scientific, legal, and social aspects of our changing world. The total Earth system provides the environment which ensures the quality of life. The International Geosphere-Biosphere Program

(continued on page 2)

HIGHLIGHTS

The **Committee on Data for Science and Technology (CODATA)** was established in 1966 by the International Council of Scientific Unions.

Working on an interdisciplinary basis, CODATA seeks to improve the quality, reliability, processing, management, and accessibility of data of importance to science and technology.

(continued from page 1)

(IGBP) is investigating the scientific aspects of all the complex interactions taking place and has defined broad research strategies requiring large data sets, with which CODATA is closely associated. This program will cover the following themes:

- human dimensions of global change (pollution, toxicology, genotoxicity...)
- data systems in biodiversity (review, state-of-the-art, network and communications)
- biogeochemical cycles (carbon cycle, nitrogen cycle)
- regional environmental change (Alps, coastal zones, regional climatic modelling)
- international global atmospheric chemistry: problems and data
- global change phenomena in terrestrial ecosystems (CGTE): existing and anticipated data (prediction scales, needs)
- sustainable ecomanagement, industrial and natural pollution control
- impediments to environmental data access [industrial proprietary data (sharing *vs.* non-sharing), cost, copyright, political obstacles]

This symposium will benefit from World Data Centers (WDC) cooperation.

Who Should Attend CODATA '94?

Scientists and industrial experts in the fields of materials and thermodynamics, specialists in environmental and ecotoxicity issues, computer scientists, scientific and technical database producers and network managers, compilers of data sets, and developers of data systems—all will find this conference valuable.

Computer Aided Data Systems and Communication will bring together experts dealing with large data systems, complex data handling, networking, and cross-discipline communication. The goal will be to improve access to, and utilization of, different types of scientific and technical databases. The focus will be on the following themes:

- bioinformatics infrastructure (including biotechnology)
- application of artificial intelligence to data management and exploitation (metadata, neural networks, knowledge extraction, data pollution, validation, and classification)
- knowledge and intelligent data banks
- management of large/complex databases (visualization, browsing, communication, organization)
- public and/or proprietary knowledge sharing strategies: world, regional and local wide area projects

CODATA participates on various levels, in the interpretation and modelling processes provided by quantitative and qualitative contributions of data, information, and knowledge collected either experimentally or created through computer simulation. CODATA naturally takes part in multidisciplinary and wide ranging problem analysis through its experience in computerized data treatment.

Its main activities involve:

- preparation of key data sets for which consistent international use is desirable
- coordination of multinational projects involving data collection and management
- promotion of new methods for data validation and prediction
- knowledge sharing and information exchange
- artificial intelligence in scientific and technical information systems
- large database management and broad network communication
- initiation of long-term projects and catalysis of new data activities through Task Groups, biennial Conferences, training courses, a *Newsletter* and specialized monographs.

CALL FOR PAPERS

All major themes of the Conference will be covered by invited lectures and contributions in both plenary sessions and parallel workshops. Interaction and discussion on complex problems will be handled in the plenary sessions; specialized topics will be dealt with in the parallel sessions.

A title and an abstract of about 500 words should be submitted as soon as possible to the Program Committee Chairman: Prof. Jacques-Emile Dubois, ITODYS, Université Paris 7, 1 rue Guy de la Brosse, 75005 Paris, France. Fax: +33 1 4288 1466. Electronic mail: via Internet to: codata@paris7.jussieu.fr

Abstracts are required by 1 October 1993. Authors will be notified before 15 January 1994 concerning the acceptance of their contribution.

Finally--A Scientific Spelling Checker!

Favorite word processing programs can now be spell-checked against technical/scientific words against agreed nomenclature, trade and trivial names, elemental symbols, surnames (for checking bibliographies), common acronyms, etc., etc. This approximately doubles the size—and usefulness—of the typical WP dictionaries.

Merged with regular dictionaries, SciWords provides an enhanced pace of spell checking of sci-tech writing by eliminating *word-not-found* messages for chemicals, agro- and biochemical names and more from *Aacaptar* to *Zyzum* in physics, too.

At about ($50 + 5) per copy it makes good sense in a scientific technical environment. Available from Pool, Heller, and Milne, Inc. 9520 Linden Avenue, Bethesda, MD 20814 (Tel. 1-301-493-6595; FAX 1-301-897-3487) or from Learned Information Ltd., Woodside, Hinksey Hall, Oxford, U.K. OX1 5AU. (Tel. 44 (0) 865-730275; FAX 44 (0) 865-736354).

On Numeric/Factual Materials and Chemical Databases

Providing chemists and other scientists and engineers with direct access to quantitative performance and properties data has taken on increased importance in recent years, and at its April, 1992, ACS Spring meeting, the Chemical and Information (CINF) Division focused a full day ACS/CODATA Symposium on that subject. Sponsorship of the sessions was shared by CODATA with J. G. (Gil) Kaufman and Dr. Gordon Wood, Secretary-General of CODATA, serving as co-chairs of the meeting. Papers from the sessions were published in the January/February 1993 (Volume 33) issue of the American Chemical Society *Journal of Chemical Information and Computer Science.*

The eleven papers that constituted the proceedings of those sessions provided a good overview of both the scope and importance of the activity and of the progress that has been made in the past ten years since the Fairfield Glade Conference* brought the need into sharp focus. CODATA was a sponsor of that original meeting, and the American Chemical Society has joined with the National Materials 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 its "proceedings" to light.

In the initial paper, Kaufman set the stage for what followed by describing the critical features of numeric data, emphasizing their focus and complexities relative to the needs of users, and by illustrating why such resources are expensive to provide. The complexity of numeric data was further highlighted in Dr. Jack Westbrook's paper on the technology and difficulties of data capture from tables and figures when converting hardcopy handbooks and publications to electronic databases.

Drs. Charles Gragg and John R. Rumble followed with accounts of the role of data collection, storage, evaluation, and interchange standards in this activity; the leadership role of ASTM Committee E49 on Computerization of Materials and Chemical Property Data was described, highlighting both the accomplishments and the challenges that remain. Dr. Anthony Barrett, then Chairman of CODATA's Task Group on Materials Database Management, provided a sense of the socio-economic implications of numeric data and, in a classic example, illustrated the importance of quality and reliability in numeric data.

One of the key factors associated with the identification of substances in databases, both for storage and for search and retrieval, is the Chemical Abstracts Service (CAS) Registry Number (RN). Dr. Charles Moulton's paper catalogued a number of enhancements made to that numbering system in recent years to increase its precision and usefulness for various materials systems.

Drs. Gordon Wood and C. Y. Ho described the responsibilities and technical approach to managing important data centers: the Canada 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 were illustrated. Another approach to providing direct access to materials data for scientists and engineers was reported in the Drago/Kaufman paper about online networking of materials and chemical databases.

The diversity of types and media for numeric/factual data was depicted by Dr. Andreas Barth's paper on an online database covering spectral and chemical structure data and by Dr. James T Staley's paper on a personal computer disk 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 Data (ASTM STPs 1017 and 1106), to CODATA Conference proceedings (e.g., "Data for Discovery", Begell House, publisher, 1992) and to various issues of the CODATA Bulletin (notably No. 69).

It is anticipated that the success of this joint symposium will encourage similar efforts in other disciplines represented in CODATA.

--J. G. Kaufman and G. H. Wood

**Computerized Materials Data Systesm, The Proceedings of a Workshop at Fairfield Glade, Tennessee*, November 7-11, 1982; Office of Standard Reference Data, National Bureau of Standards, 1983.

CODATA Calendar

1993

October

6-8 First International Symposium on Computerization and Use of Materials Property Data. Gaithersburg, MD

8-9 CODATA Task Group on Materials Property Data Managaement. Gaithersburg, MD

December

Task Group on Geothermodynamic Tables. New York (Tentative)

1994

March

17 CODATA Officer's Meeting. Paris, France

18-19 CODATA Executive Committee Meeting. Paris, France

September

18-22 International CODATA Conference. Chambéry, France

23-24 CODATA General Assembly. Chambéry, France

1996

International CODATA Conference. Tsukuba, Japan

CODATA General Assembly. Tsukuba, Japan

Computerization and Use of Materials Property Data

The Fourth International Symposium on Computerization and Use of Materials Property Data, co-sponsored by ASTM Committee E-49 on Computerization of Material and Chemical Property Data and the U.S. National Institute of Standards and Technology (NIST) will be held October 6-8, 1993, at NIST in Gaithersburg, Maryland.

Much progress has been made on the drafting of standards supporting computerization of materials property data. The principal goal of this activity has been the proper selection and usage of materials, with application of computing technology as the means to this goal. Many standards have now been approved for publication, and more are under development. The committee has more than 220 members; 20% of the membership is from outside the United States.

The purpose of the symposium is to increase the value, as seen by end users, of materials data incorporated within computerized materials information systems, which will lead to greater reliability, acceptability, and usage of these systems.

Contributions with an international perspective on any of these themes will be presented:

- End user participation in the development of computerized materials information system
- Materials-oriented applications of advanced computing technologies, such as expert systems, object-oriented databases, hypermedia, and neural networks
- Effect of user interface design upon the acceptability of computerized materials data
- Integration of computerized materials data into design and manufacturing
- Data analysis for quality and reliability
- Economic impact of computerized materials information systems in actual use
- Legal liability of developers and end users of computerized materials information systems.

More information is available from the symposium chairmen: Charles Sturrock (Tel: 301/975-6027; Fax: 301/926-7975) and Edwin Begley (Tel: 301/975-6118; Fax: 301/990-8729), National Institute of Standards and Technology, B254 Materials Building, Gaithersburg, MD 20899. (Abstracts by 4th August.)

Chemical Thermodynamic Tables

Inadvertently omitted from the list of Task Groups and Commissions 1992-94 in last quarter's Newsletter their raison d'etre is shown here:

Chemical Thermodynamic Tables. This group performs a valuable service to the scientific community as their critical evaluations and compilations save workers many hours of literature searching and prevent the errors that might occur if each worker tried to evaluate the data themselves or, worse yet, used them without thorough verification. They are in the last stages of completing their book on elemental Fe and the Fe^{++}/Fe^{+++} couple but which will also include chapters related to the wustite phase and the thermal functions of key iron compounds.

Errata

The editors apologize to Gordon H. Wood, CODATA's Secretary General for having misspelled his name in the preceding issue.

New Books

The Chemical Thermodynamics of Actinide Elements and Compounds, The Actinide Aqueous Inorganic Complexes, Part 12, by J. Fuger, I. L. Khodakovsky, E. I. Sergeyeva, V. A. Medvedev, and J. D. Navratil. (a)

International Thermodynamic Tables of the Fluid State, Volume 12, Methanol, edited by K. M. de Reuck. (b)

Strategic Issues for Electronic Records Management: Towards Open Systems Interconnection, ACCIS Working Group on Electronic Records Management Issues and Standards (WG/ERMIS). (c)

Using Online Scientific and Engineering Databases, by H. Bjelland. (d)

An Introductory Guide to Scientific Visualization, by R. A. Earnshaw and N. Wiseman. (e)

The Data Handbook: A Guide to Understanding the Organization and Visualization of Technical Data, by B. Fortner. (f)

(a) 1992. International Atomic Energy Agency, Vienna.

(b) 1992. Blackwell Scientific Publications Ltd., Osney Mead, Oxford OX2 03L, UK. 250 pp, illustrated, Approx. £32 / $58, ISBN 0-632-03455-6.

(c) The report contains: an overview of electronic records management and OSI; a detailed review of the current st of OSI standards relevant to electronic records management; a summary of the interrelationships of electronic records management and the standards, and a summary of the information support requirements. 126 pp, $32, from UN Sales Sections in Geneva and New York, UN Sales No. GV.E.92.0.13, ISBN 91-1-100374-1).

(d) 1992. McGraw-Hill, New York. 213 pp. $26.95 *pb*, ISBN 0-8306-3056-2.

(e) 1992. Springer-Verlag, New York. 156 pp. $49 *hc*, ISBN 0-387-54664-2.

(f) 1992. Spyglass, Champaign, Illinois. 229 pp. $39.95 *pb*, ISBN not stated.

Building Databases for Mechanical Test Data

A "Symposium on Laboratory Recording and Building Computerzied Data Bases for Mechanical Test Data," sponsored by ASTM standards-writing Committees E-28 on Mechanical Testing and E-49 on Computerization of Material and Chemical Property Data will be held November 16-17, 1994, in Phoenix, Arizona, in conjunction with the committees' November 14-16 standards development meetings.

This symposium will serve as a forum for those seeking to solve problems both in entering data from many diverse sources and sharing data with users, and to share information where solutions have been found and developed.

Papers approved by the peer-review process by June 1, 1994 will be published in a Special Technical Publication (STP).

More information is available from the Symposium Co-chairmen, John M. "Tim" Hold, Alpha Consultants & Engineering, 3838 Burnaby Drive, Pittsburgh, PA 15235-5028 and Harold Mindlin, CINDAS, Purdue University, 2595 Yeager Road, West Lafayette, IN 47906-1398.

Task Group on Materials Database Management

Materials Database Newsletter **April 1993, Number 18**

NETWORKS

In cooperation with the Copper Development Association, MPD Network has now added the **COPPERDATA** file to the sources available via the MPD Network on STN International. COPPERDATA includes industry standards and properties for all commercial copper-base alloys available in the USA. The PDA Engineering version of **MIL-HDBK-5** is to become the master for the MPD Network online version of MH5, due to be made available over the next 12 - 18 months, and the PDA database on polymer-matrix composites, PMC-90, will also be added to the network.

STANDARDS

Several new standards from ASTM Committee E-49 applicable to computerized material and chemical property databases have been approved by ASTM, including:

E 1309-92 - *Guide for the Identification of Polymers (excludes Thermoset Elastomers) in Computerized Material Property Databases*

E 1484-92 - *Guide for Formatting and Use of Material and Chemical Data and Database Quality Indicators*

E 1485-92 - *Guide for the Development of Material and Chemical Property Database Descriptions*

A new standard on methods of evaluation of materials and chemical property data is now in development.

For purchase of these standards and further information about the activities of ASTM E-49 on the Computerization of Materials and Chemical Property Data, contact: Teresa Cendrowska, ASTM Staff Manager, 1916 Race Street, Philadelphia, PA 19103-1187, USA.

RECENT PUBLICATIONS

ERA Technology has published a report entitled *Survey and Critical Evaluation of Materials Information Systems for Design and Research.* The 230-page report includes sections on types of computerized materials information systems; use of information on materials; general access requirements; summary of available materials information systems; an extensive review of principal information systems; market aspects, standards and new developments. ERA Report 92-0357R is priced at £150 (£135 for members) plus postage. CONTACT: ERA Technology, Cleeve Road, Leatherhead, Surrey, England KT22 7SA. Telephone: 0371 374151; FAX: 0372 377927.

CALENDAR	
	3-5 May 1993: Atlanta, GA, USA Spring Meeting of ASTM Committee E49 on **COMPUTERIZATION OF MATERIAL AND CHEMICAL PROPERTY DATA.** 5-6 May 1993: Atlanta, GA, USA First International Symposium on **COMPUTERIZED CHEMICAL DATA STANDARDS.** (This event will begin immediately following the close of the regular E49 committee meeting.) 6-8 October 1993: Gaithersburg, MD, USA Fourth International Symposium on **COMPUTERIZATION AND USE OF MATERIALS PROPERTY DATA.** Other E49 Committee meeting dates are as follows: May 1994: Montreal, Canada Nov. 1994: Phoenix, AZ, USA May 1995: Columbus, OH, USA (including Second Computerized Chemical Data Standards Symposium Fall 1995: Tokyo, Japan, Fifth Symposium on Computerization and Use of Materials Property Data

EDITOR: W. G. Jackson, The Institute of Materials, 1 Carlton House Terrace, London, UK, SW1Y 5DB. Tel +44 71 839 4071; FAX +44 71 839 2289; Tlx 8814813. There are no restrictions on the reproduction and distribution of the contents of this Newsletter.

Editor : Edgar F. Westrum, Jr.
Department of Chemistry, University of Michigan,
Ann Arbor, MI 48109
Telephone: (313) 764-7357 / Telex: 8102236056
FAX: 1-313-747-4865

Associate Editor: Phyllis Glaeser
CODATA Secretariat, 51 Blvd. de Montmorency,
75016 Paris, France
Telephone: 33 1 45250496 / Telex: 645554 F
FAX: +33 1 42881466 / Cables: ICSU Paris 016

Published four times per year (February, May, August, November)

Assistance in dissemination provided by National Committees

INTERNATIONAL COUNCIL OF SCIENTIFIC UNIONS
COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY

CODATA / NEWSLETTER

CODATA, 51 bd. de Montmorency, 75016 Paris