

HIGHLIGHTS

1. CODATA Proceedings Volumes
2. CODATA (Tsukuba) Conference Key Lectures Highlights
3. Data Quality
4. Books/Databases
5. CODATA Calendar

Data and Knowledge in a Changing World: A New CODATA Series

The two books in the series *Data and Knowledge in a Changing World* were inspired by the Symposium on Communications and Computer Aided Systems held at the 14th International CODATA Conference in September 1994 in Chambéry, France and edited by J.-E. Dubois and N. Gershon. Volume I entitled *Information Revolution: Impact on Science and Technology* is the first comprehensive, single volume book of issues concerning the challenges and the vital impact of the information revolution (including the Internet and the World Wide Web) on science and technology. Its topics concern the impact of the information revolution on science and technology:

- dramatic improvement in sharing of data and information among scientists and engineers around the world,
- collaboration (on-line and off-line) of scientists and engineers separated by distance
- availability of visual tools and methods to view, understand, search, and share information contained in data
- improvements in data and information browsing, search and access, and
- new ways of publishing scientific and technological data and information.

These changes have dramatically modified the way research and development in science and technology are being carried out. However, to facilitate this information flow nationally and internationally, the science and technology communities need to develop and put in place new standards and policies and resolve some legal issues.; The central theme of the book is the combination of a survey of several fundamental aspects of this information revolution with an evaluation of issues involved in circulating scientific and technological data. Besides policy and standards, the important issues and challenges include the maintenance of certain cultural disparities in the era of world-wide information systems, encouraging everybody to provide as well as to use information, and the changing face of collaboration in science and technology. The papers address the topics associated here with the information revolution both at the conceptual and application levels. Concrete problems in the fields of textual numerical and image databases are presented, while the complex variety of concepts and standards needed for efficient worldwide networks and intelligent access tools are demonstrated in the bioinformatics case study. However, many of the challenges and difficulties facing the information revolution and its successful implementation in science and technology will have to be dealt with and resolved in the years to come; they are essential elements of the emerging Knowledge and Information Society of the Future.

The second volume, *Modeling Complex Data for Creating Information* was similarly inspired. Whereas the first volume dealt with the numerous challenges facing the information revolution, especially its communication aspects, this one provides an insight into the recent tools provided by computer science for handling the complex aspects of scientific and technological data. It is concerned with real and virtual objects often involved with data handling processes encountered frequently in modeling of physical phenomena and systems behavior. Topics concerning modeling complex data for creating information include:

- Object oriented approach for structuring data and knowledge

- Fractal modeling and shape and surface processing
- Symmetry applications for molecular data

The choice of these topics reflects recent developments in information systems technologies. ; One example is object oriented technology. Recently, research, development, and applications have been using object-oriented modeling for computer handling of data and data management. Object oriented technology offers increasingly easy-to-use software applications and operating systems. Another example of recent development is the ability to handle imprecision and uncertainty in information systems. These systems can accommodate queries which in many scientific and technological disciplines are not precisely defined. These concepts are important for both science and technology where the norm is statistically defined and where fluctuations from the norm are expected.

A third example is the ability to handle, manipulate, and visualize physical 3D objects in real time. 3D representations through both classical and fractal methods are essential for our visual approach to reality. It is also a powerful concept for handling various overlapping data and information that can be mapped in a particular 3D space. Concepts and theories are still among the essential and basic challenges facing our Information Society. But, any knowledge emerging from raw data often involves the handling of complex data and the elaboration of new interpretations. This book discusses such processes as well as symmetry problems and fractal modeling. (see also CODATA Books, #4)

15th International CODATA Conference-Scientific Program

KEYNOTE LECTURE

M. ITO, President of the Science Council of Japan, "Recent advances and future perspectives of brain sciences"

PLENARY LECTURES

Network and Quality Control J. HIRAISHI, Agency of Industrial Science and Technology, Tokyo, Japan, "Trends and perspectives of research and development information"

Geosciences and Environment F. WEBSTER, College of Marine Studies, U. Delaware, USA, "Data issues in the geosciences and the environment"

Data Access Y. ARSKI, VINITI, Moscow, Russia

Special Lecture on Earthquakes Y. FUKAO, Earthquake Research Inst. U. Tokyo, Japan

--Plus five other Plenary Lectures TBA--

KOTANI MEMORIAL

R.-S. CHEN, Dept. of Protein Engineering, Inst. of Biophysics, Academia Sinica, Beijing, China

CLOSING SESSION

J.-E. DUBOIS, ITODYS, Paris, France A. S. KOLASKAR Dept. of Zoology, U. of Poona, Pune, India

INVITED LECTURES

Data Handling R. S. HELLER, USDA, Beltsville, MD, USA, "A proposal for publishing on the Internet"

Quality Control J.-L. DELCROIX, Laboratoire Physique des Plasmas, Orsay, France E. FLUCK, Gmelin Inst. für Anorganische Chemie der Max-Planck-Gesellschaft, Frankfurt, German SCHULTE-HILLES, Kšln, German K. NAWA, Niigata U., Niigata, Japan

Data Access K. KATO, Japan

Physics and Chemistry H. TAWARA, National Inst. for Fusion Science, Nagoya, Japan, "Data and Planning Center, NIFS—An activity as a data center" A. LI-SCHOLZ, Dept. of Physics, State U. of New York at Albany, NY, USA, "Dissemination of atomic and nuclear data—its past, present and future" J. F. DOUCET, U. of Paris, Paris, France, "Topology, fragments, and neural network in property prediction" Y. WOLMAN, Dept. of Organic Chemistry, The Hebrew U. of Jerusalem, Jerusalem, Israel, "Numerical data on the Internet: Collection, publishing, dissemination, and retrieval" K. VARMUZA, Dept. of Chemometrics, Technical U. of Vienna, Vienna, Austria, "Cluster analysis and selection of chemical structural data"

Medical Biology S. FUCHAROEN, Dept. of Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol U., Bangkok, Thailand, "Data analysis of thalassemia and related hemoglobinopathies" K. IMAI, Dept. of Physicochemical Physiology, Medical School, Osaka U., Suita, Japan, "Mutant hemoglobin databases" A. HAYASHI, Dept. of Chemistry, Kinki U., Higashi-Osaka, Japan, "Construction and application of lipid data."

Biological Macromolecules A. M. LESK, U. Cambridge, MRC Centre, UK W. C. BARKER, NBRF, Georgetown U., Washington, DC, USA, "Protein database" R. J. SIMPSON, Ludurig Inst. for Cancer Res. & the Walter & Eliza Hall Inst. of Medical Res., Melbourne, Australia

Materials F. A. KUZNETSOV, Inst. of Inorganic Chemistry, Russian Academy of Sciences, Novosibirsk, Russia

Biology D. BIGWOOD, USDA, Beltsville, MD, USA, "The USDA plant genome information system" T. SASAKI, STAFF Inst., Tsukuba, Japan, "Rice genome" N. HUANG, IRRI, Manila, Philippines H. W. MEWES, MIPS, Max-Planck-Inst. für Biologie, Martinsried, Germany, "Yeast genome" S. TABATA, Kazusa DNA Res. Inst., Kisarazu, Chiba, Japan, "Cyanobacterial genome"

Biodiversity and Systematics K. MORIWAKI, The Graduate U. for Adv. Studies, Hayama, Kanagawa, Japan, "Genetic diversity of mouse subspecies and its implications in the biomedical sciences" K. IWATUKI, Dept. Science, Hosei U., Tokyo, Japan F. A. BISBY, ILDIS Coordinating Centre, U. K. S. SUGAWARA, RIKEN, Wako, Saitama, Japan

CODATA 96 - Scientific Meetings

- PL1 Network Access
- PL2 Network, Quality Control
- PL3 Computer Technology
- PL4 Data Engineering (Sessions for Physics and Chemistry, Materials, Biology, and Genome)
- PL5 Geoscience and Environment
- PL6 Data Access (Sessions for Materials, Physics and Chemistry, and Biological Macromolecular)
- PL7 Education and Training (Sessions for Medical Biology, Physics and Chemistry, Biodiversity, and Systematics)
- PL8 Earthquake
- Plus Electronic Publication, Species 2000

Satellite Meeting Schedule

Expanded Task Group Meeting, "Survey of Data Sources in Asian-Oceanic Countries," organized by M. Tasumi* and S. Iwata; Thursday, 26 Sept.-Sunday, 29 Sept., Sanjo-Kaikan, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo

Task Group Meeting, "Biological Macromolecules," organized by T. Kunisawa, Y. Sakaki, Y. Tateno and A. Tsugita, Sanjo Kaikan; Saturday, 28 Sept. (9:00-18:00) - Sunday, 29 Sept. (9:00-12:00), The University

of Tokyo, Hongo, Bunkyo-ku, Tokyo

"Electronic Publishing," organized by John R. Rodgers, E. Fluck and S. Iwata; Tuesday, 1 Oct. (10:00-18:00), The Auditorium of Agency of Industrial Science & Technology, Tsukuba

IUBS-CODATA-IUMS, "Species 2000: Indexing the World's Known Species," organized by Frank A. Bisby and H. Sugawara; Thursday, 3 Oct. (13:00-17:00), The Auditorium of Agency of Industrial Science & Technology, Tsukuba

*underlined names are contact persons

Social Events

Welcome Buffet-Sunday Evening, September 29, Daitchi Hotel Conference Banquet

Tuesday Evening, October 1, at Satsui-toi

Technical Tours

- o A-1) Visit to National Institute of Materials and Chemical Research and Mitsubishi Chemical Corporation
- o A-2) Visit to National Research Institute for Metals National Institute for Research in Inorganic Materials or Sumitomo Chemical Company
- o A-3) Visit to Electrotechnical Laboratory, Texas Instruments or NEC
- o A-4) Visit to National Institute for Resources and Environment or the National Institute for Agro-Environmental Sciences
- o A-5) Visit to Tsukuba Life Science Center of RIKEN, Institutes of Agriculture, Forestry and Fisheries or University of Tsukuba
- o A-6) Visit to National Laboratory of High Energy Physics or University of Tsukuba

Database Quality: NIST Approach and Internet Influence - A Preview

Numeric data collections in science and engineering demand special attention to quality. Data evaluation, by definition, is an assessment of the quality of data. The Standard Reference Data Program (SRDP) at the National Institute of Standards and Technology (NIST) has long maintained a program of data evaluation, and when computer databases began to be developed, the maintenance of quality received new attention. After building and distributing more than 60 S&T databases, NIST SRDP has developed quality control procedures that it now routinely applies to all its database projects. The Internet explosion has added a new dimension, and procedures adequate for PC-based databases must be modified for online systems. We can separate the quality concerns into two parts: first, the assessment of data quality; and second, ensuring that data recommended by evaluators get into the hands of users as expected.

The evaluation of scientific data has been well described and can be examined from three viewpoints: how well is the data generation described, do the data follow the known physical laws, and do the data compare with other measurements or calculations of the same phenomena. The description of data generation is crucial. The identification and control of all relevant independent variables must be addressed and demonstrated. For mature areas such as thermodynamics and atomic physics, many measurement techniques are well characterized. Here the adherence to physical laws and intercomparisons predominate. For areas in which behavior is not well understood, such as corrosion, the physical laws are not known, and data from different experiments are not usually comparable. Consequently, documentation of control of the experimental conditions is very important. The result of the evaluation effort is a set of recommended or evaluated data, the existence of which represents value added to the original research results. Today, computer databases are becoming a primary distribution mechanism, but it is a nontrivial task to make sure that what users find when using those databases is what the data evaluator intended.

To that end, NIST SRDP has established a careful program of quality control to ensure NIST standard references databases are of the highest quality. The informal database review process developed with issuance of the first ten or so databases has evolved into a formal process with assigned responsibilities and careful checks and balances. The need for quality is reinforced by the fact that NIST standard reference databases are sold, and customers expect the same level of quality from NIST products as they do from other professional software. The purpose of the review is threefold: to ensure accuracy, correctness, and reliability.

The Internet is adding a new dimension to the quality program. Databases on the Internet have different styles and interfaces from PC-based databases. While database navigation may be similar, its implementation is considerably different. Furthermore, the Internet facilitates having systems of databases. NIST SRDP is in the process of changing both the planning and the review process to reflect these changes. New style guides will be developed, users' guides will be replaced by online help. Maintaining data integrity throughout the Internet linking process is being investigated. --Dr. John Rumble

CODATA Calendar 1996

February

- 19-20 Geothermodynamic Data Task Group (partial). Paris, France

March

- 17 CODATA Officer's Meeting. Paris, France
- 17 CODATA Publication Advisory Board Meeting. Paris, France
- 18-19 CODATA Executive Committee Meeting. Paris, France

April

- 11-12 CODATA Task Group on Materials Database Management. Paris, France

September

- 26-29 CODATA Task Group on the Survey of Data Sources in Asian-Oceanic Countries. Tokyo University Sanjo-Kaikan, Tokyo, Japan [Contact person: Prof. M. Tasumi (tasumi@tansei.cc.u-tokyo.ac.jp) Fax +81-3-38142627]
- 28-29 CODATA Task Group on Biological Macromolecules. Tokyo University, Sanjo-Kaikan, Tokyo, Japan [Contact person: Dr. T. Kunisawa (kunisawa@jpnst31.bitnet) Fax +81-475- 221544]
- 29 -Oct 3 International CODATA Conference. Tsukuba, Japan

October

- 1 CODATA Working Group on Electronic Publishing (1000 to 1800), Tsukuba, Japan [Contact person: Prof. S. Iwata (iwata@race.u-tokyo.ac.jp) Fax +81-3-34670648]
- 3 CODATA Commission on Standardized Terminology for Access to Biological Data Banks (1300-1800). Tsukuba, Japan [Contact person: H. Sugawara (sugawara@viola.riken.go.jp) Fax +81-484-624618]
- 4-5 CODATA General Assembly, Tsukuba, Japan

CODATA Books

In the series "Data and Knowledge in a Changing World," to appear:

Vol. I. Information Revolution: Impact on Science and Technology. J.-E. Dubois and N. Gershon, [a]

Vol. II. Modeling Complex Data for Creating Information. J.-E. Dubois and N. Gershon. [b]

Books and Databases

Chemical Thermodynamics of Americium. Edited by R. J. Silva, G. Bidoglio, M. H. Rand, P. B. Robouch, H. Wanner, and I. Puigdomenech. [c]

The 1996 GAPHYOR UPDATE editon (GAPHYOR UPDATE 95). [d]

COACH (COmputer Aided CHEmistry) (Database on Diskette). [e]

ISO-14000. A Guide to New Environmental Management Standards. T. Tibor with I. Feldman. [f]

Appraising the Records of Modern Science and Technology: A Guide. J. K. Haas et al. [g]

Thermodynamic Properties of Minerals and Related Substances at 298.15 K and 1 Bar (105 Pascals) Pressure and Higher Temperatures. (U.S. Geological Survey Bulletin 2131). R. S. Robie and B. S. Hemingway. [h]

Footnotes

[a,b] See Reviews on page 1. Expected February 1996. Each ca. 300 pp. Softbound. Available from Springer-Verlag, New York and Paris.

[c] Chemical Thermodynamics Series Volume 2; a critical review of the chemical thermodynamic data of those elements of particular importance in the safety assessment modeling of high-level radioactive waste storage and disposal facilities. This is a review of experimental data reported in the literature for americium. On a few occasions, where no data existed, comparisons and estimates were made based on experimental data on analog lanthanide elements. Moreover, it does contain an Appendix on "Chemical Thermodynamics of Uranium" written by I. Grenthe, M. C. A. Sandino, I Puigdomenech, and M. H. Rand which encompasses a critique and corrections to the earlier NEA-TDB review, the first volume in the series. 1995, 368 pages, ISBN 0-444-82281-X Hardbound, Dfl. 375.00, US\$234.50. North-Holland; P.O. Box 211, 1000 AE Amsterdam, The Netherlands; P. O. Box 945, Madison Square Station, New York, NY 10159-0945, USA.

[d] Published by the beginning of each year, includes all data collected during the previous year. The 1996 edition includes 25,745 references. Note: the actual version 1.2 is still a b-version, with some possible problems. More information is available on the WWW server, <http://gaphyor.lpgp.u-psud.fr>. 850 FF, plus taxes. Denis Humbert et al., Centre de DonnŽes GAPHYOR, 91405 Orsay.

[e] In several mouse clicks, thanks to an intuitive graphic presentation, access is given to thermodynamic properties of more than 4000 chemical substances and their reactions. Hypertext access immediately provides advice, instructions, definitions, and theoretical considerations, as well as power and user friendliness. Single and group licenses are available. Thermodata, B. P. 66, 38402 St. Martin d'Herès CEDEX, France (1995).

[f] ISBN 0-7863-0523-1. xvi + 238 pp. Softcover. Price \$35.00. Irwin Professional Publishing, 1333 Burr Ridge Parkway, Burr Ridge, IL 60521.

[g] "All other archival activities hinge on appraisal. This guide is intended to assist in appraising the records of modern science and technology. The twentieth century has been significantly influenced and shaped by science and technology; but the record of post-World War II science and technology is vast, and not all can be preserved. In order to identify and preserve an appropriate and useful record for the future, archivists need to appraise knowledge." 1985. 76 pp. Softcover. Distributed by the Society of American Archivists, 600 South Federal, Suite 504, Chicago, IL 60605.

[h] Selected values for S° , V° , $D_{\text{for}}H^\circ$, $D_{\text{for}}G^\circ$ on elements, 55 sulfides, 93 oxides and hydroxides, 141 silicates, and 154 other minerals and related substances at 298.15 K. Where available, appropriate functions are tabulated at 100 K intervals for temperatures up to 1800 K. Available from U. S. Geological Survey, Information Services, Box 25286, Federal Center, Denver, CO 80225. iv + 464 pp. \$29.00

Editors

Editor: Edgar F. Westrum, Jr. Department of Chemistry, University of Michigan Ann Arbor, MI 48109-1055
Telephone: (313) 764-7357 / Telex: 8102236056 FAX: +1-313-747-4865 E-mail: westrum@chem.lsa.umich.edu

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 E-mail:
codata@paris7.jussieu.fr

Published four times per year (February, May, August, November) Assistance in dissemination provided by
National Committees

CODATA home pages: <http://www.cisti.nrc.ca/codata/welcome.html>