INTERNATIONAL COUNCIL OF SCIENTIFIC UNIONS
ISSN 0538-69
COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY

# 67 CODATA NEWSLETTER

JULY 1994

## HIGHLIGHTS

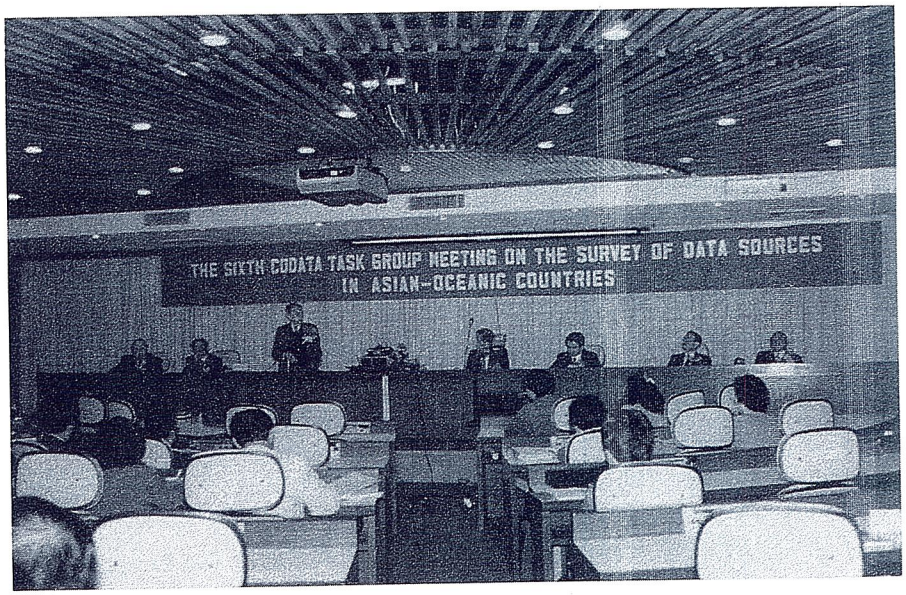


The **Committee on Data for Science and Technology (CODATA)** was established in 1966 by the International Council of Scientific Unions.

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

## The CODATA Task Group Data Sources in Asian-Oceanic Countries Meets in Taipei

The Sixth Meeting of the CODATA Task Group on the Survey of Data Sources in Asian-Oceanic Countries was held at the Activity Center, Academia Sinica, Taipei, Taiwan, March 9-10, 1994. About 80 participants got together from Australia, China (Academy of Sciences, Beijing), China (Academy in Taipei), India, Indonesia, Japan, Pakistan, and the Philippines. In addition, two honored guests from CODATA were




also present: Professor David Abir (President) and Dr. Gordon H. Wood (Secretary General). Unfortunately we had no participants from Korea, Malaysia, and Thailand.

After a warm welcome address from Professor Tung-bin Lo, (Vice President of Academia Sinica), Professor Abir, Professor Mitsuo Tasumi (Chairman of the Task Group), and Dr. Jen-Leih Wu (Chairman of the Organizing Committee) made responses. Thirty-eight presentations from the participants followed through the next morning.

*(continued on page 3)*

## CODATA National Committees

***Interesting Data Developments.*** A search of the 1994 reports of the National Committees for innovative data activities of their own construct or underway in their own countries has revealed the following activities.

### Australia

***Metadata system.*** NRIC has established the National Directory of Australian Resources (NDAR), which currently contains descriptions of a range of land resource and geoscience datasets. Information on soils, forestry, and environmental datasets will be added shortly.

A special feature of NDAR is that it is spatially searchable; a dataset's description is retrieved whenever the search window intersects its defined coverage. Another important feature is that it can store information on datasets described in different ways (*e.g.*, databases and maps) and retrieve all relevant records in response to a single search request.

***Special reference should be made to the impact that Internet and WorldWideWeb have made in scientific communication.***

NDAR is a spatially-searchable directory containing comprehensive descriptions of the content, spatial coverage and subject matter of natural-resource and related datasets. Common directory software developed by NRIC is now installed in a number of Federal and state agencies. Links are proposed to NASA's Master Directory, and NRIC is keen to develop links with other national and international directory systems.

### Canada

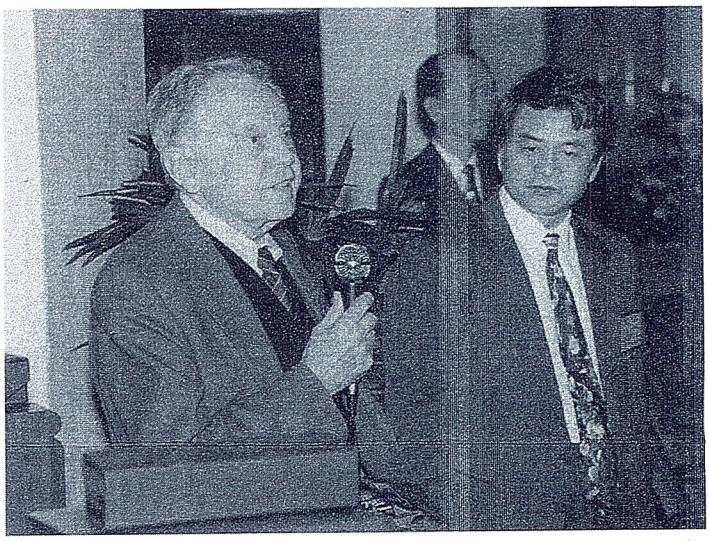
***Chemical Thermodynamics.*** F*A*C*T is a Canadian thermochemical database system which contains thermodynamic properties on over 4000 inorganic stoichiometric compounds (5000 phases) including aqueous and gaseous ions. The public system is accessed via X.25 networks with host computers at McGill University, École Polytechnique de Montréal and CISTI (NRC, Ottawa).

### China

***Data Highway.*** To make full utilization of database resources by users all over the country, CODATA-China started to organize information services of science and technology via telecommunication network. Some institutes or research groups were selected as service units in which information resources of the same discipline are highly concentrated. In the service unit a network

*(continuued on page 4)*

## Asian-Oceanic TG



*CODATA's President David Abir addresses the TG Meeting; Dr. Ken-Leih Wu listens.*

## Asian-Oceanic TG

*(continued from page 1)*

About 24 persons, including Professor Abir and Dr. Wood, attended the discussion session on the future activities of the Task Group, chaired by Professor Tasumi, on March 11. Professor Tasumi discussed his views on the application for renewal of the CODATA Task Group, and the participants reacted to this presentation.

The Chairman also mentioned fruits in the form of publications of our past Task Group activities, noting particularly the forthcoming *Presentations* of the '94 and '92 meetings and the *Directory*.

Regarding proposed activities, the Chairman explained setting up three subgroups, namely i) animal viruses, ii) microbes, and iii) materials. Additional proposals were sought—one on "fish" was proposed.

It was a commonly appreciated that all four might not be encouraged by the General Assembly because of the heavy burdens and severe financial restrictions that might be encountered. Some discussions to establish priorities on them followed. The Chairman was, however, reluctant to decide the order of priorities by vote and put his decision off until the Subgroup leaders present him more exact and detailed proposal documents.

The leaders of the four proposed Subgroups were designated as:

- Survey of Databases on Animal Viruses:
  Professor A.S. Kolaskar
- Survey of Databases on Microbes:
  Dr. Hideaki Sugawara
- Survey of Databases on Fish:
  Professor Kwang-tsao Shao
- Survey of Databases on Materials:
  Professor Shuichi Iwata

Each leader was requested to submit a revised text of the corresponding proposal to the Chairman, who would revise the whole text and send it to Paris for the Executive Committee Meeting.

The possible changes of the TG membership were discussed. Dr. Rudy H. Tan, from the Philippines, and Professor Kazuo Takayanagi, from Japan, were replaced by Mr. Glenn L. Sipin and Professor Yukikazu Itikawa. New additions were Ms. Rukasih Dardjat, from Indonesia, Professor Zafar H. Zaidi, from Pakistan, and Professor A.S. Kolaskar, from India. The possible retirement of Professor Tadashi Makita and the possible addition of Professor Kwan-tsao Shao, from China were considered.

Discussion then followed about how to treat those graphical data which are difficult to handle.

The next Task Group Meeting is scheduled to be held in Tokyo or in Tsukuba, Japan, in 1996, adjacent to the 15th CODATA International Conference.

Appreciation was expressed for the excellence of the arrangements for the Meeting and the heartfelt hospitality of Professor Wu and his staff on the Organizing Committee.

*—prepared by Makato Kizawa*

**CODATA Task Group Meeting on the Survey of DATA Sources in Asian-Oceanic Countries. Academia Sinica, Taipei, Taiwan.**

## CODATA National Committees

*(continued from page 2)*

environment of computer telecommunication for databases is set up using the PSTN (Public Switched Telephone Network) as the major means of access at present. The information services of science and technology include data retrieval, problem-oriented consulting, development and offering of specialized software, speciality computations, etc.

For the first phase (1994-1995) two service units were selected to give trial services. One of these in the Laboratory of Computer Chemistry of the Institute of Chemical Metallurgy, Chinese Academy of Sciences, provides service in chemistry and chemical engineering, and another in Tsinghua University provides service in materials. An experiment of online retrieval to the Advanced Ceramics database of Tsinghua University from the Institute of Chemical Metallurgy was carried out.

***Information Network.*** Since the late seventies, 19 databases of *ca.* 6.3 gigabytes of data on science and technology have been built in relevant institutes of the CAS, mainly on engineering chemistry, mass spectra, carbon-13 NMR, rare earth materials, microbial resources, infrared spectra, astronomy, Chinese medicinal plants, economic plant resources, natural resources, materials corrosion, x-ray photoelectron spectroscopy. And in the Zhongguancun area where the institutes of the CAS are concentrated, an optical fiber computer network called NCFC (National Computing and Networking Facility of China) was completed in 1992. This connects the campus networks in Tsinghau University and Peking University and CAS network (CASnet). The CASnet consists of 19 Local Area Networks of 20 Institutes. In 1993 the *Computer Network Information Center (CNIC) of the Chinese Academy Sciences (CAS)* was organized to implement development, retrieval, integration, consulting and training of resources of information and networks on science, technology and education. According to plan, most of the 19 databases will be transferred to the CNIC host to offer services.

## China (Academy located in Taipei)

***Rummaging/Compiling.*** As the principal makers of the world's inexpensive personal computers, we have inexplicably underutilized them. Managing data are what computers do best. Fortunately, this survey uncovers some encouraging trends. Some scientists discovered that those fading graphs, idle numbers, and handbooks left to them by scientists of yesteryear can become treasure troves. That is, you use computers to build databases out of them, and again rely on the computers to make use of the databases. At a time when online services provided by vendors and CD-ROM networks are commonly available, this new turn seems paradoxical. However, scientists and engineers alike have also come to appreciate the power of computer-aided simulation, modeling, and forecasting by using well-made databases.

***Vanishing Gene Pools.*** Even before the global attention is focussed on the earth and its ecosystem, we have become acutely aware of our diminishing gene pools. Hence, the protection of existing life forms to preserve biodiversity has taken on a new order of importance. The establishment of gene banks, ranging from seeds to cells, is but a natural course to take. Taiwan's scientists in the life sciences and agriculture are racing with time to preserve something of the past and the present. Several databases in agriculture and the life sciences reflect the deep concern among the scientist working in these fields.

## CODATA Calendar

1996

| | |
|---|---|
| ------ | International CODATA Conference. Tsukuba, Japan |
| ------ | CODATA General Assembly. Tsukuba, Japan |
| ------ | CODATA Task Group on the Survey of Data Sources in Oceanic Countries. Japan |

## Japan

***Biodiversity.*** The Task Group of Data Sources in Asian-Oceanic Countries decided to initiate a survey of databases on microbes in Asian-Oceanic Countries as one of its subjects. The sub-group will be headed by Dr. H. Sugawara, WFCC World Data Center on Microorganisms (WDCM). WDCM is the component of World Federation of Culture Collections and actually managed by RIKEN. WDCM published the fourth edition of the World Directory of Collections of Cultures of Microorganisms (bacteria, fungi, and yeasts) in 1993 and distributed copies to 481 collections registered in WDCM. The Task Group's subjects and the activities of WDCM will be one of the cores of information management of biodiversity projects.

***Environmental Protection.*** The Task Group meeting of Data Sources in Asian-Oceanic Countries on March 10-12, 1994, decided to set up a subgroup to make a survey of databases on environmental protection in Asian-Oceanic Countries. Plans for the survey and establishment of databases relating to environmental protection in Asian-Oceanic Countries are proposed by Dr. S. Iwata. Funding for this project would be sought inside Japan.

## South Africa

***Psychology Databases.*** The databases relating to psychology in this country are to be found in, for example, the NAVO database which is managed by the Centre for Science Development of the HSRC. This database contains information on both ongoing and completed research projects of senior students in psychology and by members of the social sciences community in this country. The latter category would include members of staff of psychology departments at universities, the staff of the HSRC, projects undertaken by various civil service departments, and so on. In general, NAVO serves as a useful tool to ensure that scarce financial resources are not squandered by duplicating research, and it also serves as a means to enable psychologists in South Africa to form informal networks.

## United States

***Combining Diverse Environmental Data.*** The report, *Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data*, was authored by a subcommittee of the USNC/CODATA, chaired by Dr. G. Bruce Wiersma of the Univer-

*(continued on page 6)*

## CODATA Books

CODATA Proceedings Series

**Vol. 1. Data Sources in Asian-Oceanic Countries; DSAO, Taipei, 1994,** edited by Jen-Leih Wu, Yaruo Hu, and E. F. Westrum, Jr. (a,b)

**Vol. 2. New Data Challenges in Our Information Age; Beijing, 1992 CODATA Conference,** edited by Phyllis Sloane Glaeser and Michael T. I. Millward (a,c)

**Vol. 3. Database Developments in Asian-Oceanic Countries; DSAO, Beijing, 1992,** edited by Yaruo Hu and Edgar F. Westrum, Jr. (a,d)

CODATA Monograph Series

**Vol. 1. Crop Modeling and Related Environmental Data: A Focus on Applications for Arid and Semiarid Regions in Developing Countries,** edited by Paul Uhlir and G. C. Carter (a,e)

**Vol. 2. CODATA Directory of Data Sources in Asian-Oceanic Countries,** edited by Yaruo Hu and E. F. Westrum, Jr. (a,f)

## Books

**Activity Coefficients at Infinite Dilution, Vol. IX, Part 3,** by J. Gmehling, J. Menke, and M. Schiller (g)

**Activity Coefficients at Infinite Dilution, $C_{10}$ - $C_{36}$ with $O_2S$ and $H_2O$ - Supplement, Vol. IX, Part 4,** by J. Gmehling, J. Menke and M. Schiller (h)

**Manual on the Building of Materials Databases,** edited by Crystal H. Newton (i)

**ASTM Standards on the Building of Materials Databases** (j)

**Navigating the Networks, Proceedings of the 1994 Mid-Year Meeting of the American Society for Information Science,** edited by D. L. Anderson, T. J. Galvin, and Mark D. Giguere (k)

**Guide to Database Distribution, Second Edition,** by Joseph Bremmer (l)

## Databases

**DIPPR Version 9.2.** (m)

**Hazardous Materials, Cancer Therapeutic Agents, New Pharmacologically Active Compounds, Organic Photochemistry,** and **Peptides and Amino Acids** (n)

**NIST Critical Stability Constants of Metal Complexes** (o)

---

***Footnotes***

(a) Published by CODATA. Volumes in both the Proceedings Series and the Monograph Series may be purchased either from the CODATA Secretariat or the Ann Arbor, Michigan "outpost." (See both addresses on page 6 of this Newsletter.)

(b) Published 1994, 318 pp., Softbound $39 (US) or equivalent in Fr. francs, postage included.

(c) Published 1994, 672 pp., Soft-/Hardbound $50/$80 (US) or equivalent in Fr. francs, postage included.

(d) Published 1994, 176 pp., Softbound $33 (US) or equivalent in Fr. francs, postage included..

(e) Published 1994, 256 pp., Soft-/Hardbound $34/$69 (US) or equivalent in Fr. francs, postage included.

(f) Published 1994, 208 pp., Softbound $33 (US) or equivalent in Fr. francs, postage included

(g) Published 1994, 505 pp., DECHEMA Chemistry Data Series, Vol. IX, Part 3, Frankfurt am Main, Germany. Price for Part 3 and 4 together: 690 DM. ISBN: 3-926959-46-0.

(h) Published 1994, 568 pp., DECHEMA Chemistry Data Series, Vol. IX, Part 4, Frankfurt am Main, Germany. Price for Part 3 and 4 together: 690 DM. ISBN: 3-926959-47-9.

(i) Published 1993, 105 pp., Soft cover. $45 list. ASTM, 1916 Race Street, Philadelphia, PA 19103-1187.

(j) Published 1993, 112 pp., Soft cover. $43 list. ASTM, 1916 Race Street, Philadelphia, PA 19103-1187.

(k) Published 1994, 255 pp., Soft cover. $29.95. From Learned Information, 143 Old Marlton Pike, Medford, NJ 08055-8750. ISBN: 0-938734-85-7.

(l) Published 1994. $175 (including shipping and handling). Order by prepayment to NFAIS, 1518 Walnut Street, Suite 307, Philadelphia, PA 19102.

(m) Late release of the Design Institute for Physical Property Data Pure Component Data Compilation. The program, for IBM PCs and compatibles, provides data on 26 constant and 13 temperature-independent properties of 1,405 chemicals. $2035 single user, Technical Database Services.

(n) Chemical Abstracts Service, $995 per title. Five new titles in the CASurveyor line of CD-ROM products. Each CD-ROM contains up to three years of bibliographic, abstract, and index information from the Chemical Abstracts database. CASurveyor is compatible with Windows and Macintosh computers.

(o) Provides information on reaction rates of ligands with ions. The database covers interactions in aqueous systems of nearly 4,000 organic and inorganic ligands with protons and various metal ions. The program runs on DOS-based PCs. $240. National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899.

## ASTM Standards on Electronic Transfer of Environmental Lab Data

Newly formed ASTM Task Group E50.04.02 on Environmental Laboratory Export Format will develop standards for the electronic transmission of environmental data between laboratories and their users. According to Leah Reed, chairman of task group which is part of Committee E-50 on Environmental Assessment, laboratories analyze millions of environmental samples each year. By using computers to transfer the results of these analyses, time is saved, paper work is reduced and transcription errors are virtually eliminated.

The task group will write standards for a single format for the laboratories to produce and the receiver to deal with, even though different analytical methodologies may be used. The task group wants a format that can be used for data on organics, inorganics, and for radiochemical data.

## CODATA Task Group on Fundamental Constants

The TG meeting was held 2 July 1994 at Boulder, CO, USA, the day after the completion of the 1994 Conference on Precision Electromagnetic Measurements. In attendance were the following Task Group members: B. N. Taylor, *NIST*, USA (Task Group Chairman); F. Cabiati, *IEN*, Italy; E. R. Cohen, *Rockwell International*, USA; T. Endo, *ETL*, Japan; V. Kose, *PTB*, Germany; Liu Ruimin, *NIM, PRC*; B. W. Petley, *NPL*, UK; T. J. Quinn, *BIPM*; B. M. Wood, *NRC*, Canada.

A number of experiments of critical importance to the next least-squares adjustment of the constants are still underway. The results of these experiments, some of which were expected to be reported at CPEM 94, will not be available until the end of 1994 or in early 1955. They include two values of the Planck constant from the NPL and the NIST moving-coil watt balances; and two values of the fine-structure constant from the CSIRO/NML (Australia) and the NIST calculable capacitor-quantum Hall effect experiments.

It was concluded that the *a-priori* assigned uncertainty of each potential input datum should be carefully reviewed and increased if the physics of the experiment seems to warrant such an expansion, and possibly increased if it (*i.e.*, the datum) overwhelmingly dominates the adjustment.

It was also concluded that the tables of recommended values in the 1987 *Reviews of Modern Physics* article describing the 1986 adjustment are quite satisfactory and should not be altered significantly.

## CODATA National Committees

*(continued from page 4)*

***From the United States report —***

sity of Maine. It was published in the summer of 1994 and is available free of charge from the USNC/CODATA office.

***Transborder Flow of S&T Data.*** The primary task of the Committee for the next two years will be conducting and publishing its new study, *Bits of Power: Issues in the Transborder Flow of Scientific and Technical Data.* The study will focus on data in electronic forms, a topic of increasing complexity and importance in scientific research and international collaboration. The resulting report will characterize the technical, legal, economic, and policy issues that have an influence—whether favorably or negatively—on access to S&T data by the scientific community. Special attention will be given to the specific conditions inherent in the transborder transfers of electronic S&T data among the academic, governmental, and private sectors. The report will then identify and describe those barriers that have the most adverse impact in discipline areas within CODATA's purview—the physical, astronomical, biological, and geological sciences—and across those disciplines, using representative examples. Finally, it will identify trends likely to have significant discipline-specific and interdisciplinary influence on the use of S&T data.

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

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

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

Assistance in dissemination provided by National Committees

COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY

CODATA / NEWSLETTER

CODATA, 51 bd. de Montmorency, 75016 Paris