



# 56 CODATA / NEWSLETTER

MAY 1991

## HIGHLIGHTS

Beijing CODATA Conference	1,(2,3)
Biological Macromolecules	6
Biological Terminology	4,(5)
CODATA Calendar	5
Executive Committee	5,(6)

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.

## "New Data Challenges in Our Information Age"

**The 13th International CODATA Conference, CODATA '92, will be held on 19-22 October 1992 in Beijing, China.**

The Conference will be organized around sessions in the area of scientific and technical data challenges. It will provide an appropriate framework for discussion of the following major themes:

- Global change
- Artificial intelligence: data and knowledge synergy
- New computer-aided archaeological methods
- Biology and agriculture
- Materials and industrial data
- Geological and space sciences
- Chemistry and physics
- Computer networking
- Programming tools
- Legal and social aspects of data

Certain more specific themes will also be featured.

Nobel Laureate (1986) Professor Yuan T. Lee of the University of California at Berkeley has agreed to present the Conference Keynote lecture.

Plenary and invited lectures on energy sources, materials, food composition, crop germplasm, and chemistry will be presented by Chinese scientists. For example, a plenary lecture on Peking Man—a resident of Zhoukoudian 200 000 to 500 000 years ago is scheduled. Afterwards, a scientific site visit to Zhoukoudian will provide an enhanced appreciation of human evolutionary development. Other prominent lecturers will be featured in the August Newsletter.

*Continued on p. 2)*



## CODATA '92

(continued from p. 1)

In addition, three sessions in parallel with appropriate CODATA sessions will be offered by the World Data Centers. They are entitled:

Paleoclimate data  
Environmental data rescue, and  
Access to environmental data.

### Call for Papers

All major themes of the Conference will be covered by invited papers. In addition, contributed papers will be presented, either in oral or poster form. Specialists in the fields of data collection, analysis, evaluation, and organization of data sets and data systems applications are invited to submit contributions on subjects within the scope of the Conference. To facilitate discussion and interchange of ideas, workshops will be organized from the poster sessions in which groups of related posters will be discussed with an appointed chairman after the poster sessions.

A title and abstract of about 500 words should be submitted as soon as possible to either of the Program Committee Chairmen:

Prof. J. E. DUBOIS	Prof. XU Zhihong, Director
CODATA Secretariat	Institute of Chemical Metallurgy
51, bd. de Montmorency	Chinese Academy of Sciences
75016 Paris, France	PO Box 353
	100080 Beijing, China

Electronic mail to Paris may be sent on OMNET: CODATA.PARIS; BTGOLD: 10075:DB10010; or DIALCOM: 145:CDT0010.

An abstract will be required by 30 September 1991. Authors will be notified about the acceptance of their contribution in November 1991 and will receive instructions on providing the full text at that time. The first announcement may be sent on or about July 10, 1991, the second one about October 10, 1991.

### Who Should Attend CODATA '92

Practitioners in all scientific disciplines, especially those in the physical, biological, geological, and astronomical sciences concerned with the management of quantitative data resulting from experimental measurements or observations, will find this Conference beneficial.

### Exhibits

Display space will be available for organizations which desire to inform the participants about their equipment, services, and books relevant to CODATA interests.

### Database Demonstrations

Ample space will be available for scientific and technical database producers and developers to conduct demonstrations of their equipment and/or services. Software and hardware requirements should be requested from the Business Agent in advance.

### Participation Expenses

Until July 19, 1992 the conference registration fee will be (US)\$230. Thereafter it will be (US)\$280. Conference participants (but not accompanying persons) will enjoy a reception and a banquet without charge. For students a special (US)\$60 fee is available. The per room per night rates at the chosen hotel are (US)\$80. Special student housing will be available for (US)\$25. China International Air Line may offer approximately 30% reduction of round trip prices to China from Europe, USA, etc. for Conference participants.

### Social and Cultural Activities

Social and cultural activities planned for participants and their accompanying guests attending CODATA '92 include the following beside the reception and banquet:

- Tours of China's unique ancient sites and exquisite gardens; Palace Museum, the biggest palace of the emperors of Ming and Qing Dynasties; the Ming Tombs, Beihai Park; etc.
- Visits to the world-renowned Great Wall, Summer Palace, a former Royal Garden over 800 years old, unique ancient Chinese architecture of the Temple of Heaven with a history of more than 500 years, the Echo Wall, the Circular Mound Altar, etc.

Ample time has been set aside for guests to discover Beijing's charm, culture, and art, and to enjoy different cuisine offered by well-known Chinese chefs, independently of the organized program.

Beijing—located at the northern rim of the North China Plain, bordered by the Loess Plateau in the West, and by the Inner Mongolia Plateau in the North—is at the junction of flatlands, mountain ranges, and highlands. It is a city of scenic beauty with a population of over 10 million. The nation's most important scientific research centres—77 institutions of higher learning, the Chinese Academy of Sciences, as well as the Chinese Academy of Social Sciences together with many other research organizations—are located in Beijing. Its charm lies in being an ancient capital with many splendid cultural sites, beautiful gardens, and landscapes. Over a hundred cultural, historical sites and places of interest make it a tourist centre as well as a cultural and scientific exchange centre. With many famous chefs, the hotels in Beijing provide specialities from Sichuan, Shangdong, Guangdong and Huaiyang, as well as Western food. Famous restaurants: *FANSHAN RESTAURANT* serves imperial court cuisine, *FENGZEYUAN RESTAURANT* offers Shangdong cuisine, and *QUANJUDE ROAST DUCK RESTAURANT* is well known both at home and abroad for Peking Duck.

### Post-Conference Tours

Tours can be arranged including two days visiting in Beijing and then tours to Xi'an, one of the important cradles of Chinese civilization, and famous ancient capital of 11 dynasties of China; to Guilin, a charming city with magnificent peaks, caves, and streams, which is claimed to be "second to none in scenery"; and to Shanghai, one of the biggest commercial metropolises of China, and one of the world's largest seaports. It is the seat of Fu Dan and Tongji Universities, other colleges, institutes,

(continued on p. 3)



## CODATA '92

(continued from p.2)

museums, libraries, and parks; to Hangzhou, a famous city with beautiful scenic spots—West Lake, Zhongshan Park, Lingyin Temple, Dragon Well Spring, Jade Spring; etc. Tours of other cities may also be arranged on request.

### Local Organizing Committee

Prof. XU Zhihong, Chairman, *Institute of Chemical Metallurgy, CAS*

Assoc. Prof. HU Yaru, Vice-Chairman, *Computing Center, CAS*

Mrs. ZHANG Shizhen, Vice-Chairman, *Computing Center, CAS*

Mr. XIAO Yun, Business Agent, *Computing Center, CAS*  
Mr. ZHU Shixue, *Institute of Mathematics, CAS*  
Prof. JIANG Chusheng, *Chinese Ministry of Chemical Industry*  
Mr. QIU Xibai, *Chinese Chemical Society*

Further Conference scientific and practical information can be obtained from the Conference Business Agent:

Mr. XIAO Yun, Business Agent, CODATA '92  
No. 4, Nan 4th Street  
Zhongguancun, Haidian District  
Beijing 100080, CHINA  
Tel: +86-1-2561410 FAX: +86-1-2562485

Expected participants are urged to clip (or Xerox) this form and send it to the Conference Business Agent [Mr. XIAO Yun (address above)] via air mail to arrive soon—and preferably before 10 September 1991. (Those desiring to present papers or posters should also provide title and abstract to one of the Program Committee Co-Chairmen. See **Call for Papers**.)

I am interested in **CODATA '92**, the 13th International CODATA Conference entitled "New Data Challenges in Our Information Age" to be held in Beijing, China, from 19-22 October 1991.

☐ Please send the preliminary program when available.

☐ I intend to contribute a paper ☐ orally ☐ at a poster session

Major Theme: \_\_\_\_\_

Tentative Title: \_\_\_\_\_

Titles and abstracts (about 500 words) shall be sent to a Program Committee Co-Chairman as soon as possible (See Call for Papers.)

☐ Please send information on arranging for exhibits of \_\_\_\_\_ equipment/services \_\_\_\_\_ books.

☐ Please send information on providing database demonstrations.

☐ I will probably be accompanied by \_\_\_\_\_ persons.

☐ I (with \_\_\_\_\_ other persons) plan to take the following local tours:

☐ Great Wall (near Beijing), October 18

☐ Two day's tour around Beijing, including Great Wall, Palace Museum, and Temple of Heaven, October 23-24

☐ We are interested also in post-Conference tours starting October 25 to ☐ Xi'an, ☐ Guilin, ☐ Shanghai (Hangzhou). Please send information.

My name, address, etc. are: (please print)

Name: \_\_\_\_\_ Title: \_\_\_\_\_

Affiliation: \_\_\_\_\_

Street Address: \_\_\_\_\_

State/Province: \_\_\_\_\_ ZIP/Postal Code: \_\_\_\_\_

Country: \_\_\_\_\_ Tel. No. \_\_\_\_\_

FAX No.: \_\_\_\_\_



# Problems of Access to Standardized Biological Terminology

The CODATA Commission on Standardized Terminology for Access to Biological Data held a Workshop in Nancy, France on 14-16 May 1991 to formulate a plan for improving access to standardized terminology for biological database producers and users. The Workshop, jointly sponsored by the U.S. National Center for Biotechnology Information and the Commission of the European Communities DG XII, was attended by representatives of the Biological Unions of the International Council of Scientific Unions (ICSU), producers of bibliographic and factual databases, and professional terminologists. This combination of participants, coming from disparate subdisciplines of biological and information science, provided an excellent blend of appropriate talents to address the multi-faceted problems of terminological standardization.

The essence of the Workshop was summarized by a participant: "I have learned a great deal from this Workshop.....I know now that I have been working too much in isolation and must begin to see more of the bigger picture."

One of the primary goals of the Commission, reemphasized during the Nancy Workshop, is to enhance the level of consciousness within the biological community of the need to communicate across disciplines. A major benefit of today's computer technology is that it provides the means to integrate data in ways which will lead to new scientific insights. Artificial intelligence, innovative programming, massive data storage capabilities, and vastly improved communication technology, will inevitably draw diverse data sources together. If data are to be integrated, exchanged, and searched efficiently, the intellectual input to make this possible must come from biologists now.

Although problems surrounding the "standardization" of nomenclature and terminology have been with us for centuries, information technology demands that we take a new look at these problems and devise new methods of solving them which take full advantage of today's technological tools.

The Workshop program was divided into three segments. The first consisted of a number of formal presentations that provided a background and overview of the perceived problems in interdisciplinary access to biological terminology. Dr. Andrzej Elzanowski, of the Max-Planck-Institut für Biochemie, articulately summarized the problems faced by database producers in "translating" the nomenclature and terminology used by authors writing articles for scientific journals into some type of "standard" that can be used for consistency in retrieval of identical concepts. It became obvious during the course of the Workshop that authors, editors, publishers, and database producers are all facing a similar situation, i.e., the lack of clear guidelines on nomenclature and taxonomy for organisms and the terminology to describe their characteristics.

Representatives of the biounions then had the opportunity in a series of round table discussions to present the "state of the art" within each Union regarding nomenclature and terminology standards. Broad topics covered by the Union representatives included botanical and zoological nomenclature, taxonomy, biochemistry, microbiology, pharmacology, physiology, nutrition, food science, and clinical medicine. During discussions it became apparent that the nomenclature and terminology com-

mittees of the ICSU Unions face just as many problems in providing standards as database producers and users face in locating standardized terminology for biological concepts. Their ability to provide wide access to standardized terminology in formats and electronic media desired by database producers is limited by the fact that virtually all of their work is done on a "volunteer basis" and is primarily designed for intradisciplinary use. Additional resources would be required to expand the work of the biounion nomenclature and terminology committees.

During the third segment, participants heard terminological specialists make the group aware of existing standards for terminology. There are general principles which must be applied when developing terminological databases, regardless of the scope and content. It was agreed that the Commission would take cognizance of these principles and would work with the specialists to educate the biologists on the ICSU committees in the implementation of these principles. Several documents recognized by the International Standards Organization were recommended for use in the educational campaign.

The dialog engendered during the Workshop raised the awareness of the participants that the present interdisciplinary nature of many scientific activities leads to a greater need for exchange of information within the component parts. The impact of multinational projects, e.g., HUGO, similarly imposes added demands for clarity and standardization of expression. The integration of international, interdisciplinary databases will require some precision in defining terminology for uniform interpretation of scientific principles.

It was decided that the first step in providing wider access to standardized terminologic references is the expansion of the effort initiated by the U.S. National Library of Medicine in establishing a "Nomenclature File" in its *Directory of Biotechnology Information Resources*. All agreed that the NLM file is a useful beginning, but that a broader international inventory of terminological resources and their relationships to one another is required. A Steering Committee composed of Commission members will be established to design procedures for developing this international terminological inventory.

The ultimate goal of the Commission would be to catalyze the development of an international "Term Bank." This, of course, would have to be developed in modules and would necessitate preliminary studies to determine feasibility and user requirements. The Commission would also seek the cooperation of other international organizations such as the International Council on Scientific and Technical Information (ICSTI) and the International Federation of Scientific Editors (IFSE); both groups were represented at the Nancy Workshop. The enormous effort required to make such a Term Bank available to the international scientific community was recognized. While the computer and communication technology is available to link subsets of such a database, issues such as copyright, cost recovery, coordination, updating responsibilities, and funding were recognized to be potential barriers to the successful accomplishment of the goal.

Dr. Leslie Sobin, the representative from the International Union Against Cancer (UICC) aptly set to verse the precautions which



## Standard Bio-Terminology

(continued from p. 4)

must be taken by the Commission in approaching a project of this dimension:

*"When you're looking for the answer  
how to classify all cancer,  
proteins, microbes, fish and succulent legumes,  
You must know a little Latin,  
tell a round fish from a flat one,  
and have memory with lots and lots of room.*

*But, before we start alinking"  
we should sit back and be thinking"  
on our methods, clientele and on our goal,  
Lest we make a mammoth bank  
rarely used and rarely thanked  
just consuming funds and efforts: A BLACK HOLE.*

--Leslie H. Sobin

The CODATA Commission will certainly keep these words of wisdom in the forefront as it launches its campaign to improve access to standardized biological terminology.

Further information on the CODATA Commission on Standardized Terminology for Access to Biological Data can be obtained from the CODATA Secretariat, 51 bd. de Montmorency, 75016 Paris.

--Dr. Lois Blaine

## CODATA Executive Committee Meeting

The CODATA Executive Committee met 25-27 March 1991 in Paris and several CODATA groups held meetings more or less coordinated with the Executive Committee. In particular the Executive Committee as a whole undertook a long-range planning session to work out major directives for modifications of CODATA endeavors and the new standing Publications Committee under the chairmanship of Professor E. Fluck considered new directions to achieve the goals of CODATA publications.



CODATA 1991 Executive Committee Meeting, Paris Secretariat. Clockwise from nearest table corner: R. SINDING LARSEN, R. ECKERMANN, E. FLUCK (invitee), A. TSUGITA, R. SIMPSON, J. CREASE, J.E. DUBOIS, G. WOOD, D. ABIR, D. LIDE, P. GLAESER, F. KUZNETSOV, L. GURVICH, J.H. WESTBROOK, G.C. CARTER. Not shown: E.F. WESTRUM, JR. and Z. XU (invitees).

**Access to Data.** Dr. Lide reported on a survey of producers-, vendors-, and users of data within the scientific community. The

## CODATA Calendar

1991

June

----- Chemical Thermodynamics (Part A) Task Group. Aachen, Germany

21-24 Chemical Thermodynamics (Part B) Task Group. Moscow, U.S.S.R.

22-28 Intl. Symposium on Chemistry and Chemical Thermodynamics. Moscow, U.S.S.R.

August

8-10 CODATA IGBP Commission. Paris, France

September

9-11 3rd Intl. Symposium on Computerization and Materials Data. Cambridge, U.K.

1992

March

16-18 Executive Committee Meeting. Paris, France

April

13-15 Materials Regularities Workshop. Como, Italy.

October

19-22 International CODATA Conference. Beijing, China

23-24 CODATA General Assembly. Beijing, China

*ad hoc* Study Group which conducted the survey included Dr. G. Wood, Dr. David Watson, and Dr. D. Confland. It received about 70 responses to the 200 questionnaires disseminated in seeking responses to the questions concerning the magnitude of four potential types of barriers inhibiting ready access to scientific data, i.e.: political, legal, economic, and technical barriers. The five questions involved more particularly:

- (1) transmission across national boundaries
- (2) differences in pricing and subsidy problems
- (3) the cost problem for academic scientists
- (4) barriers resulting from attempted protection of intellectual properties
- (5) the particular problems of developing countries.

There seemed to be little concern about political barriers. The application of copyright law to electronic databases seemed a concern to many respondents but this was not judged to be a significant barrier. On the economic aspects, although there was wide diversity—and even contradictory opinions—in the responses, three themes predominated:

- the economic value of scientific and technical data (STD) is not properly appreciated
- the (generally) current practice of subsidizing PRODUCERS rather than USERS is not the most effective use of funds
- STD's are a commodity and should be subject to market forces

(continued on p. 6)



## Meeting on Biological Macromolecules

The CODATA Task Group on Biological Macromolecules convened for the first time in the United States, at Georgetown University Medical Center in Washington, DC, on March 13-16, 1991. The meeting was hosted by the Protein Identification Resource (PIR) at the National Biomedical Research Foundation. Funding for travel, lodging, and conference facility fees was provided by a grant from the National Science Foundation. The meeting began with a one-day colloquium of formal presentations by special guests around the theme "Integration of Information Concerning Biological Macromolecules."

Harold Morowitz introduced the concept of the "Matrix of Biological Knowledge" or "biomatrix." Lois Blaine explained the mission of the CODATA Commission on Standardized Terminology for Access to Biological Data. John McCarthy described new computer tools for genome research being developed at Lawrence Berkeley Medicine, particularly those of the new National Center for Biotechnology Laboratory. Dennis Benson outlined the initiatives of the National Library of Information, in providing access to biological data. Lois Hunt illustrated

the usefulness of value-added protein sequence databases. Xiaojie Xu described the Peking University Protein Molecular Modeling and Design System. Lastly, Ashok Kolaskar described the use of sequence, structure, and hybridoma data banks to predict the structure of an antigenic viral protein. About 60 people attended this colloquium.

The Task Group meeting was attended by ten Ordinary Members, nine Corresponding and Consulting Members, fourteen special guests, two representatives from the National Science Foundation, and several members of the PIR staff. This is the largest group ever to attend of meeting of the Task Group. Discussion was lively during both the formal meeting and the breaks. It appeared that a number of new collaborations will be initiated as a result of this meeting.

### Executive Committee

*(continued from p. 5)*

Proponents of the last mentioned view obviously do not recognize the existence of a problem. That such a viewpoint exists is a concern of itself in failing to appreciate the innate worth of STD to the international scientific community. Perhaps ICSU—and more particularly CODATA—needs to provide some educational "propaganda."

The stratagem of subsidizing the user rather than the producer is intriguing.

The overwhelming majority of respondents seem to feel that there should be "academic discounts" on grounds of general social benefits but noted that a distinction between pedagogical and research uses needs to be drawn. The general plight of university libraries, many of which cannot afford to purchase needed journals and books, to say nothing of electronic information services, has received much attention. However, this should not be viewed as simply a problem of access to data, but as an integral part of the much broader problem of inadequate support of basic scientific research.

Likewise, some felt that a prudently set differential might exist between national uses and those in other countries. Most respondents agreed that third-world scientists have little chance of using the new electronic information services. There are both economic and technical barriers (i.e., poor telecommunications, lack of computers, etc.). However, the information gap is perceived as only one of many problems which impede the creation of a viable R & D infrastructure in many developing countries; other priorities may be higher. More training courses and other educational activities would clearly have benefits.

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

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: 630553  
FAX: +33 1 42889431 / Cables: ICSU Paris 016

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