COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY

61 CODATA / NEWSLETTER

DECEMBER 1993

13th International CODATA Conference

Beijing, China, 19-22 October 1992

Entitled *New Data Challenges in our Information Age,* the 13th International Conference of the Committee on Data for Science and Technology was uniformly agreed to be a scientific success by the 283 participants, about half of whom came from 21 countries other than China. This biennial Conference is a key forum for specialists in many disciplines to gather, share problems, exchange ideas, and propose solutions relating to the evaluation and handling of scientific data.

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.

Russian Academician F. Kuznetsov discusses the future of electronic materials.

A novel feature at this gathering were several joint sessions with the ICSU Panel on World Data Centers who were meeting concurrently with CODATA. Presentations included 5 general plenary lectures, about 80 lectures in 21 discipline-oriented parallel sessions, and a combination of about 120 oral discussions and posters in a series of 6 topical workshops.

Reflecting the collaborative nature of the Conference, the opening Plenary speaker John Townshend of the University of Maryland outlined the data requirements and role of the International Geosphere-Biosphere Program (IGBP). As an example, he observed that those modelling global circulation or ecosystem dynamics need temporal data varying over times ranging from seconds to centuries and spatial data varying in resolution from individual trees to continents. One of the great challenges to CODATA is the integration of existing and future data sets. Even the detection of change, something a non-expert might regard as relatively straightforward, is virtually impossible if data sets gathered from a variety of sources cannot be accurately geographically registered.

To make a planned afternoon tour of the famed Peking Man Site at nearby Zhoukoudian more meaningful, Wu Rukang of the Institute of Paleoanthropology of the Chinese Academy of Sciences gave a plenary lecture on their current excavation and research activities. In the third plenary presentation, Shizuo Fujiwara of the Japanese Research Institute for Information and Knowledge elegantly summarized some of the problems inherent in East-West communication and information transfer. To illustrate, he observed that for schemes involving automation, it would be ideal if each scientific term in one language mapped into unique terms in other languages. In reality the picture is far more complex. Because a given term in one language may have several meanings depending on the scientific field wherein it is used, any automated system conveying information from one language to another must be sensitive to field-specific usage. He suggested that communication system designers would do well to examine the manner in which the three Japanese "alphabets" combine to permit varying degrees of generality and specificity in expression.

The other plenary lectures were given by Fedor Kuznetsov, Institute of Inorganic Chemistry of the Russian Academy of Sciences, and Jacques Dubois, University of Paris. Prof. Kuznetsov spoke on the future of electronic materials developments. While other applications have numerous alternative materials upon which to rely, he felt that electronic processes will be limited to silicon for the foreseeable future and outlined the features of an impressive database system that will facilitate research into all aspects of this critical material. Prof. Dubois, who also served as Co-Chairman of the Scientific Program Committee, proposed a molecular paradigm to address the complexity of the information age. Just as nature is ordered from quanta, electrons and neutrinos on one scale to atoms and molecules on the next scale to polymers, cells, organisms, and ecosystems on a third scale, so should we think of organizing and accessing data. Instead of just storing and retrieving isolated data items, the emphasis may have to shift to storing associated information such as the patterns and systematics that occur in the data.

To adequately summarize or even characterize the multitude of other presentations is clearly not possible; only a few arbitrary chosen highlights will be cited. Some idea of the scope of topics covered can be grasped by considering this partial list: data handling; social, political and legal aspects of data; IGBP interfaces; material databases; intelligent databases; genome projects; resources and environment in Asia; molecular modelling; thermodynamics; biology; spectroscopy; paleoclimatic data; large data sets; emerging methods for data handling; artificial intelligence and data exchange formats.

Paul Mezey, University of Saskatchewan, Canada, gave a particularly intriguing presentation on problems associated with the non-visual shape analysis of molecules. He reminded his audience that molecules "see" each other, not as the ball and stick models that humans visualize, but as clouds of electron density whose shape depends on the pertinent density level. Because storing this virtually infinite variety of three-dimensional shapes in a computer for chemical modelling purposes is clearly impractical, other strategies must be developed.

In a much different paper, Senliang Li *et al.* of the Scientific Database Center of the Chinese Academy of Sciences, described their database system dealing with Chinese Medicinal Plants. Drawing on sources in which their country is rich, they have compiled a database containing more than 2000 medicinal plants, 1500 medicinal chemicals, and 5000 traditional Chinese medicines. Each entry includes descriptors pertaining to its botanical, chemical, and pharmacological characteristics along with a bibliographic reference. In one sense the database system is a reflection of modern China—a blending of the new and the very old—for the system incorporates the latest in computer technology to find bibliographic citations that may be 500 to 1000 years old!

In the session on IGBP Database Interfaces, Gottfried Gabert, Germany, dealt with the interfacing of alphanumeric, graphic, and remote sensing databases by map construction for application to geosciences, resource assessment and land-use planning. The Graphics Interactive Raster Oriented System, developed at the Geological Surveys in Hannover, is an important tool for constructing hybrid digital maps from both vector (*e.g.*, Map) and raster (*e.g.*, satellite) data. Such interfacing will play a significant role in a forthcoming CODATA-related project in the Himalayas.

S. Ohsuga from Japan and Ning Zhong from China, in the session on Practical Applications of Artificial Intelligence, addressed the difficult problem of how to organize intelligent databases, showed an operational example using a relational database and contributed a methodology which should work for other databases. During discussion of the paper, the problem was raised of how knowledge derived from databases can be sheltered from classical knowledge in the sciences so that the former, with time, does not substitute the latter and thus impoverish our total knowledge.

An important aspect of this conference was the joint participation of the World Data Centers (WDC) group. In a September 1992 meeting the CODATA/FAGS/WDC agreed to closer cooperation in an ICSU *ad hoc* committee chaired by Dr. Michael A. Chinnery.

--G. H. Woods
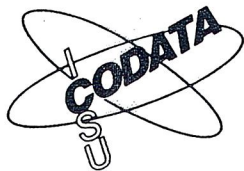
Task Group on Materials Database Management

Materials Database Newsletter   December 1992, Number 17

The Ninth meeting of the CODATA Task Group on Materials Database Management took place in Paris in September and all tasks in hand were reviewed and progressed. A second edition of the International Register of Materials Database Managers will be published early in 1993. Its scope, in terms of number and content of entries and geographical coverage, will be considerably wider than that of the first edition which was issued in 1990.

A questionnaire on the costs of database building and use has been distributed to Task Group members for their comments, and a final version of this questionnaire will be distributed widely among database managers early in 1993. The correlated results of this investigation, together with the Task Group's latest analysis of the benefits of materials databases and it ongoing study of a standardised description of databases, will be fed into the programme of a new Workshop which the Task Group is planning. This CODATA Workshop on the Costs, Benefits, and Economic Consequences of Materials Property Database is scheduled to take place early in 1994.

The Task Group has welcomed a proposal to work in concert with a new Group on Materials Database Technology, initiated by the Association for Science Cooperation in Asia (ASCA) which held a seminar on Materials Database Technology in February 1992 (see below).

NETWORKS	**MPD Network** has announced plans for the addition of two new databases. The **ASMDATA** file will include all the databases produced by ASM International as PC databases (Mat.DB) and will cover both metals and plastics. MPD Network and the Optiz Corporation, St. Petersburg, Russia, have signed an agreement to produce a new database, **OPTIMAT**, on the properties of optical materials. The database will concentrate on glass but will also include optical glue, single crystals, optical polymers, and metals.
RECENT PUBLICATIONS	**Proceedings of the ASCA Seminar on Materials Database Technology, February 5-7 1992, Tokyo, Japan**
	This seminar was held under the auspices of the Association for Science Cooperation in Asia (ASCA). The purpose was to exchange experience and opinions concerning computerized materials information systems by materials experts from ASIAN—Oceanic countries. The requirements for, and developments in, materials data and computerised materials data systems were considered in a number of country reports from Australia, Peoples Republic of China, India, Indonesia, Korea, Malaysia, The Philippines, Thailand, and Japan. The importance of international collaboration, for example through the activities of CODATA and VAMAS, was discussed. The seminar included a round table discussion leading to a number of recommendations for future actions. FURTHER INFORMATION: Dr. S. Nishijima, National Research Institute for Metals, 2-3-12, Nakameguro, Meguro-ku, Tokyo 153, JAPAN.
CALENDAR	1-3 February 1993: Frankfurt am Main, GERMANY
	CODATA-CEC Joint Workshop on **MATERIAL DATA FOR COMPUTER AIDED ENGINEERING.** CONTACT: DECHEMA, Abt. I&D, Theodor-Heuss-Allée 25, D-6000 Frankfurt/Main 15, GERMANY.

3-5 May 1993: Atlanta, GA, USA

Spring Meeting of ASTM Committee E49 on **COMPUTERIZATION OF MATERIAL AND CHEMICAL PROPERTY DATA**. CONTACT: Teresa Cendrowska, ASTM, 1916 Race Street, Philadelphia, PA 19103-1187, USA.

6-8 October 1993: Gaithersburg, MD, USA

Fourth International Symposium on **COMPUTERIZATION AND USE OF MATERIALS PROPERTY DATA**. CONTACT: ASTM, 1916 Race Street, Philadelphia, PA 19103-1187, USA.