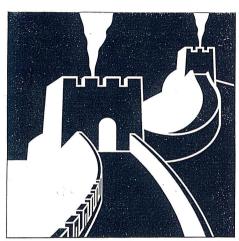


SEPTEMBER 1991



CODATA—BEIJING—1992

HIGHLIGHTS

CODATA Calendar 4
CODATA Comm. IGBM
CODATA Personalities 6
Mat'l. Database NL 7(8)

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.

Why Do We Need Relevant Food Data Interchange?

Max Feinberg*, Jayne Ireland-Ripert, Jean-Claude Favier Centre Informatique sur la Qualité des Ailments 16 rue Claude-Bernard, F-75231 Paris Cedex (France)

CODATA TASK GROUP ON A SYSTEMATIC NOMENCLATURE FOR FOODS IN NUMERICAL DATA BANKS

Foods are involved in many human activities and, although a few million people have too little or no food to eat, we can be certain that several billion meals are eaten each day. Hence, foods are involved in all major human activities:

- health and science, in the context of food correlated diseases such as cancer or diabetes, or malnutrition in developing countries;
- trade and economy, with increasing international exchange of commodities;
- regulation and politics, which depend on food hygiene, agricultural policy and trade control;
- social behavior and human sciences, as food is not only a way to survive but also a source of pleasure and involved in many social or religious habits.

Unfortunately, we are not always able to understand and compare nutritional status for different countries or people, due to the fact that there is no scientific method for describing foods. Natural language is often inadequate and even misleading to those who are not closely acquainted with the local language and culture; for instance, the wide family of sausages represents an infinite variety of meat products containing more or less beef, calf, and/or pig (and even donkey) meat with various additives such as starch, fat or chemicals according to local regulations; it is particularly difficult to differentiate such products from one country to another. The situation is further confused by homonyms, synonyms, identical brand names for different products, and culinary or technological terms [1].

Significant food description is also of increasing importance due to labeling regulations [2]. Since September 24, 1990, nutritional labelling is regulated at

(Continued on p. 2)

Food Data Interactions

(continued from p. 1)

the European level, following similar legislation in the U.S.A. and Canada [3]. A normalized method of describing foods will be necessary, as the obligation of nutrition labelling will have an effect on international trade.

Thus, there is a veritable need for an international food language.

This need really became patent when several databases on food nutrient composition were created in different countries and data interchange was attempted. In 1988, several users and compilers of nutrient databases decided to create, within CODATA, a working group in order to discuss the construction of an international food language. The goals of this new CODATA Task Group were defined as follows, according to their order of priority:

- provide an international system by which foods can be described for databases and especially for nutritional databases, a system for the description of both food items and food samples;
- design procedures for collecting information on food composition and data interchange;
- standardize methods of data processing to obtain certified values from aggregated data;
- establish guidelines for database structure.

For the members of the Task Group, access to information was considered generally more important than actual data exchange. During the first two years, three formal meetings were organized thanks to the sponsorship of the U.S. National Cancer Institute. A faceted vocabulary, called LANGUAL, was adopted as the basic framework for developing the international food language. Initially, this system was developed by the U.S. Food and Drug Administration (FDA) to provide a standardized language for describing foods and specifically for classifying food products for information retrieval [4]. Designed to serve FDA needs, the vocabulary describes those characteristics that affect the safety and/or nutritional quality of food products. As constructed, it is a thesaural system using faceted classification but does not conform to the ISO standard.

Table 1 presents an example of a standardized description of a food using LANGUAL: French bread produced in France according to French regulations. An interesting feature of LANGUAL is that the coded descriptions of American, English and German breads are significantly different from one another because these are not the same foods, either from a nutritional, technological or legal standpoint [5]. A typical description consists in a list of 11 to 23 standardized descriptors put together in the same thesaurus. Each descriptor belongs to a facet, specific to a set of characteristics useful for food identification, identified by an uppercase letter. In order to avoid imprecision and errors of translation, each descriptor has a scope note which contains a legal or technical definition. In order to enlarge or narrow retrieval equations, descriptors are structured as hierarchical trees within each facet.

TABLE 1. LANGUAL codes for the bread produced in France

| Factor | | Code | Descriptor | |
|--------|----------------------------|------|--|--|
| A | Product Type | 0178 | Bread | |
| В | Food source | 1421 | Soft wheat (Triticum aestivum) | |
| C | Part of plant/animal | 0208 | Seed or kernel, skin removed, germ removed | |
| E | Physical state or shape | 0105 | Whole shape achieved by forming, thick, 1.5-7 cm | |
| F | Extent of heat treatment | 0003 | Complete heat transformation | |
| G | Cooking method | 0005 | Baked or roasted | |
| H | Treatment applied | 0256 | Carbohydrate fermented | |
| J | Preservation | 0003 | No preservation method used | |
| K | Packing medium | 0003 | No packing medium used | |
| P | Consumer/dietary group | 0024 | Human food, no age specification, regular diet | |
| M | Container or wrapping used | 0003 | No container or wrapping | |
| N | Food contact surface | 0003 | No food contact surface present | |
| Ç | Establishment | 0002 | Retail establishment | |
| 3 | Geographical origin | 0250 | France | |
| 5 | Storage conditions | 0001 | Storage conditions unknown | |
| Γ | Production period | 0689 | Production June 1989 | |

At present, LANGUAL is translated and usable in four languages (English, French, Danish, and Hungarian). It is implemented on several computers for retrieving data in about 10 databases concerning chemical analysis, residue values of perticides, toxic elements, nutrient elements, industrial chemicals, food additives and regulatory information. Altogether, about 30,000 food products have been coded in various countries using this system.

Often terms employed for a descriptor are very closely associated with regulation, food supply and/or culture. For example, the definition of "skimmed milk" differs from one country to another as shown in Table 2. Therefore, when precise information is available, descriptors must include a precise spectrum value instead of a simple vernacular term.

TABLE 2. Example of different food regulations. The fat content of milk

| Food name | Europe [6] | U.S.A. [7] |
|---------------------------------------|---------------------------------|----------------------------|
| Whole milk Half-skimmed or low fat | \geq 3.5% fat 1.5 to 1.8% fat | ≥ 3.25% fat 0.5%, 1%, 1.5% |
| Skimmed | ≤ 0.3% fat | and 2% fat ≤ 0.5% fat |

CODATA Commission on IGBP Data — Planning Meeting

A planning meeting of the CODATA Commission on IGBP (International Geosphere-Biosphere Program) Data was held on August 8-10, 1991, in the CODATA Secretariat, Paris. Dr. Michael A. Chinnery, Commission Chairman and about a dozen participants mainly from CODATA participated.

After discussion of actions that led to its formations by the Chairman, Committee members spoke about proposed activities such as:

- data directories, standards, and documentation (Dr. Tchtiaque Rasool)
- ecological research, terrestrial ecosystems, biological diversity (*Dr. Frank Bisby*)
- an internationally sponsored workshop, funded by Unesco with Dr. G.C. Carter as convener in March 1992 (*Prof. Shelton Alexander*). This constitutes the first short term project to be undertaken.
- environmental changes in the Himalayan region in conjunction with the Center for Integrated Mountain Development (*Prof. Claude Bardinet*)
- development of "template" software to facilitate integration of local and regional global change data with global change data (*Prof. R. Sinding-Larsen*)

In the general discussion which followed, the CODATA role was further defined and plans made for identifying helpful sources. Activities relating to the human dimensions of global change were of special interest and expertise would be sought. Although the IGBP Working Group on Data and Information Systems is heavily involved in the compilation of data sets needed by IGBP core projects, they are less concerned with some important basic data management issues. These included the following issues:

- Data Rescue: Much instrumental environmental data collected during the last 100 years or so are of great importance for the study of global change. Yet in many cases these data are not being properly stored, and in some cases are in imminent danger of being lost. An international program to identify and save these data is needed.
- Data Preservation: The study of global change requires that data collected today be preserved in a form that will be useful to researchers 50, 100, or more years hence. Very few institutions throughout the world are concerned with long-term preservation of environmental data. There are no accepted international standards for the archiving of global change data. CODATA is in a good position to begin to establish such standards.
- Data Quality: It is, of course, not enough to just preserve global change data. The data must be thoroughly documented and subject to quality control procedures if they are to be worth preserving. Issues such as data validation and data continuity are very important. No international organization has so far concerned itself with the issue of data quality. This is an area where CODATA could make a significant contribution to the global change project.

• Data Access: A main requirement for global change data is that they be readily accessible to research scientists. There are many aspects to the problem of data access. Data must be easy to find, through the use of directories and catalogs. Data must be as free as possible from commercial or national restrictions to access. Analog data need to be converted into digital form for use by modern computer-oriented researchers. Cost should not be an impediment to the use of global change data. CODATA could take a leading role in promoting international access to global change data.



Figure legend: IGBP Commission Planning Meeting (l to r): J. Rose (Unesco), M. Chinnery (Chairman), F. Bisby (IUBS), C. Bardinet (France), J. Crease (CODATA), R. Tomlinson (IGU), S. Alexander (USNC-CODATA), R. Sinding-Larsen (CODATA), P. Glaeser (CODATA). Not shown: I. Rasool (IGBP).

The planning group noted that two recent documents emphasized the importance of the above four issues. The first was the recent statement of official United States policy on "Data Management for Global Change Research," issued by the US President's Advisor, D. Allan Bromley, on July 2, 1991. The second was the recent CODATA survey on "Barriers to Access to Scientific and Technical Information." In particular, the U.S. policy is a significant step forward, but these policies need to be carried into the international arena for discussion and acceptance.

Noting that some of these four issues are relevant to the activities of the ICSU Panel on World Data Centers, the ideas were presented to the panel meeting in Vienna, Austria, on August 21, 1991, and the ICSU panel is already moving to establish a task group to take up data rescue on an international scale.

Working groups in the other three data areas have been proposed and the ICSU Panel would be invited to participate. Members for the commission and its working groups were discussed and appointments will be made through CODATA's President, Prof. David Abir.

CODATA Calendar 1991 November 15-16 Task Group on Geothermodynamic Tables. Uppsala, Sweden 1992 March 2-5 International Geosphere-Biosphere Program Commission on Data. Chambery, France 9-10 Two-dimensional Electrophoresis Subgroup of CODATA TG on Biological Macromolecules. Paris, France CODATA Biological Macromolecules Task 12-14 Group. Paris, France Task Group on CODATA Referral Database. 14 Paris, France 16-18 Executive Committee Meeting. Paris, France April 13-15 Materials Regularities Workshop. Como, October 19-22 International CODATA Conference. Beijing, 23-24 CODATA General Assembly, Beijing, China



Asian-Oceanic Data Sources Task Group

Beijing Hotel. Conveniently located to Tien An Men Square, The Forbidden City, and the main shopping area on Wang Fujing Street, the 900 room Beijing Hotel China is the site for the October 1992 CODATA Conference.

Reminder: New CODATA Task Groups

The time for suggesting creation of new CODATA TASK (and/or WORKING) GROUPS is approaching. These are to be made on a form available from the CODATA Secretariat and should be submitted by December 15, 1991, to enable Executive Committee action in its spring meeting.

Free Connect Time on the MPD Network

MPD Network is an integrated numeric information service that includes data on the properties of engineering materials. An easy-to-use menu interface and an online thesaurus help searchers locate specific materials and the detailed properties that characterize those materials. Materials currently covered in the MPD databases include polymers and both ferrous and nonferrous metals; engineering ceramics and composites will be added in the near future.

Those interested in more information about this free offer on the MPD Network should contact the MPD Network, P. O. Box 02224, Columbus, Ohio 43202; phone: (614)447-3661; FAX: (614)4473713; or TELEX: 6842086.

STN International, the scientific and technical information network, is operated jointly by CAS in North American; in Europe, by FIZ Karlsruhe; and in Japan, by JICST, the Japan Information Center of Science and Technology. A network of more than 100 databases, STN International offers information on a broad range of scientific fields, including engineering, thermodynamics, materials science, physics, biotechnology, and chemistry.

CAS, in turn, offers software products to make online searching easier (STN Express), printed products to provide current awareness in specialized areas (CA Selects and Chemical Industry Notes), easy online access to numeric data (Materials Property Data Network), and special services to make the search for information and retrieval of documents quick and convenient (CAS Search Service and CAS Document Delivery Service), etc.

Food Data Interactions

(continued from p. 2)

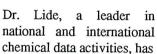
In conclusion, there is certainly a need for international collaboration in the development of an international food language.

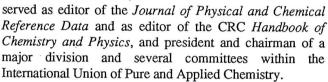
References

- * Dr. Feinberg is the TG Chairman
- [1] W. Polacchi, "Standardized food teminology: an essential element for preparing and using food consumption data on an international bases," *Food and Nutrition Bulletin*, Vol. 8, No. 2, pp. 67-68.
- [2] Commission du Codex Alimentarius, "Rapport de la vingtième session du comité du codex sur l'etiquetage des denrées alimentaires," Ottawa, 3-7 avril 1989.
- [3] EEC Regulation 90/496 of 24 September 1990.
- [4] A. McCann *et al.*, "FDA's Factored Food Vocabulary for food product description," *Perspectives in Practice*, March 1988, Vol. 88 n°3, pp. 336-341.
- [5] M. Feinberg, J. Ireland-Ripert, J.C. Favier, "LANGUAL: un langage international pour la description structurée des aliments," *Solence des Aliments*, 11 (1991), pp. 193-214.
- [6] EEC Regulation 1411/71 of 29 June 1971.
- [7] Code of Federal Regulations 21, chapter 1.

Dr. David R. Lide is the 1991 recipient of the Patterson-Crane Award for the documentation of chemistry of the Columbus and Dayton Sections of the American Chemical Society.

As Secretary General and subsequently as President of CODATA and now as Past President, he coordinated international data programs and database development in many areas of physical, biological, geological, and astronomical sciences.





He has had a long and distinguished career both as a research scientist and the leader of the Standard Reference Data Program of the US National Institute of Standards and Technology (formerly National Bureau of Standards) (1954-1988). As the director of the program, he oversaw a national effort to develop reliable databases of physical, chemical, and materials properties.

Dr. Lide is currently editor of the CRC Handbook of Chemistry and Physics. He is a Fellow of the American Physical Society and has served on many committees and Editorial Boards of the American Physical Society, the American Institute of Physics, the American Chemical Society, and other professional organizations. Previous awards include Department of Commerce Silver and Gold Medals, Samuel Wesley Stratton Award of the National Bureau of Standards, and Herman Skolnik Award of the ACS Division of Chemical Information.

Dr. Lide has authored over 100 papers on molecular structure and spectroscopy, free radicals, molecular lasers, and various aspects of scientific information.

The Award is given in honor of two former editors of *Chemical Abstracts*, Austin M. Patterson and E.J. Crane, for outstanding achievements in chemical information theory or practice. The Award was presented at a joint meeting of the ACS Columbus and Dayton Sections in Columbus, Ohio, on Tuesday, May 21, 1991.

The rules of eligibility for the award state that a nominee shall have made "contributions of national notice" to the documentation of chemistry or to chemical information theory or practice. Such contributions may be in the areas of the production of books, articles, reviews, and bibliographies; editorial work; abstracting; nomenclature; construction of indexes, codes, or methods of classification; development and use of methods for searching the literature and of mechanical or electronic aids to information storage and retrieval; chemical library work; or related activities.

Dr. Ronald L. Wigington, Director of Chemical Abstracts Service (CAS) since 1986, has been named Director of Information Technology for the American Chemical Society (ACS), the parent organization for CAS. In his newly created position, Dr. Wigington will carry out special assignments of crucial strategic importance associated with the information service activities of the Society. Dr. Wigington also served as director of Washington operations for ACS from 1984 to 1986, and director of research and development at CAS from 1968 to 1984.

This management top change follows the recent formation of the CAS Governing Board to oversee the business aspects of CAS operations. CAS has just set up a joint venture with Molecular Design that will change the character of accessing much of CAS's database. Continuing to work at CAS headquarters in Columbus, Ohio, Wigington will remain largely



associated with information service activities.

Dr. Wigington's associations with CODATA have been many, both at US National Committee level and in the international operation. He currently serves as a co-opted delegate to CODATA for the International Council of Scientific and Technical Information, as a member of the US National Committee for CODATA, and as contact for a CODATA Supporting Organization-CAS.

Professor Loren G. Hepler has worked on the CODATA Task Group on Chemical Thermodynamic Tables to establish reliable thermodynamic property values. He was honored by a 28 paper Symposium during the August 1991 Calorimetry Conference in DeKalb, Illinois. The Symposium held on the occasion of his retirement from the Department of Chemistry of the University of Alberta included his lecture on the treatment of the relaxation contribution to the heat capacities of aqueous electrolyte solutions.

(continued on p. 6)

Y.-I. Mashiko, Deceased September 1991

Yo-Ichiro Mashiko, 73, died 20 September 1991 while hospitalized in Tokyo. The funeral ceremony dedicated to him was held on Sunday afternoon, 22 September, at a Buddhist temple in Tokyo where some three hundred people attended to express their sympathies.

Dr. Mashiko will be remembered by CODATA scientists as a very active Japanese participant.

In the early 1980's he was not only the Chairman of the Japanese National Committee, but he represented Japan in the CODATA General Assembly as Delegate (1980-84). Moreover, he was also the Chairman of the Organizing Committee for the Kyoto



International CODATA Conference in 1980. He was a very productive member of CODATA's Executive Committee in the 1984-86 era.

His scientific achievements were concerned with bridging the gap between governmental scientific physico-chemical and analytical chemical practices and the rapidly developing Japanese chemical industries in the 1970's and 1980's. For many years he served in the directorship of the National Chemical Laboratory for Industry in the Shibuyaku region of Tokyo.

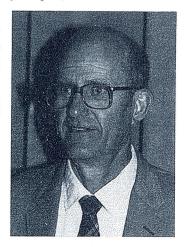
He was an avid connoisseur of Japanese cusine and delighted in introducing others to the more advanced aspects of eating in Japan. His genial influence and generous welcoming made visits to Tokyo a special treat.

He is survived by his wife, Minoru, whose address is 4-31-9 Nogata, Nakano-ku, Tokyo 165.

Loren G. Hepler

(continued from p. 5)

He has made many contributions to chemical thermodynamics covering a very wide variety of subjects including calorimetry, aqueous and nonaqueous solutions, transition metal chemistry, substituent and solvent effects related to organic acids, clay chemistry, water treatment, hydrogen sulfide chemistry, and the science and technology of oil (tar) sands and heavy oils. He has authored or co-authored 230 scientific publications. Loren



Hepler has served as co-editor of the AOSTRA Journal of Research and on the editorial boards of four other journals. He has been actively involved in the technology of oil sands and heavy oils as Chairman of Management for the AOSTRA-Alberta Research Council Oil Sands Research Project. He served on the Board of Directors of the Calorimetry Conference from 1976 to 1981, and as Chairman from 1978-1979. He has also been active in civic affairs and is a recipient of the Van Cleave Award from the University of Regina and an honorary DSc degree from the University of Lethbridge. He is a recipient of the Huffman Award, the highest honor of the Calorimetry Conference.

--Robert N. Goldberg NIST, Gaithersburg, MD

Texas A&M Database in Moscow

Kenneth N. Marsh was recently seen in Moscow, USSR, expounding the virtues of the TRC Database for Chemistry and Engineering: Vapor Pressure. The occasion was the CODATA and IUPAC sponsored Symposium on Calorimetry and Chemical Thermodynamics on the centennial of the Lugunin Thermochemical Laboratory of Moscow State University in June. (Another CODATA stalwart, Prof. Lev Gurvich, was Chairman of the International Program Committee and a member of the Organizing Committee.)



Ken Marsh presents TRC Database in Moscow.

Ken—IUPAC's delegate to CODATA—has also been named Editor of *Chemical and Engineering Data* this year.

The new, computerized vapor pressure database from the Thermodynamics Research Center makes it easy to access an extensive compilation of vapor pressure and boiling point data on organic, inorganic, and organometallic compounds. The database, now available on diskette for IBM PC-compatible computers, contains Antoine constants for more than 5700 compounds.



Task Group on Materials Database Management

Materials Database Newsletter

August 1991, Number 14

NETWORKS

MPD Network is now available worldwide on STN International. Databases are searchable by either STN command mode or via MPD Network menu-driven interface. Databases now available include: International Plastics Selector, MARTUF, Metals Datafile, PLASPEC and STEELTUF, plus greatly expanded versions of aluminum alloy files AAASD and ALFRAC. MPDSEARCH, a directory of worldwide data files, and the MPD Thesaurus are also available. New files now being added include the NIST Structural Ceramics Database, NISTCERAM, and COPPERDATA. Polymer Matrix Composites will be added in 1993. Access to MPD network is by subscription for unlimited search or by usage-based fees. CONTACT: J.G. Kaufman (Tel: +1 614 447 3706) or Bill Weida (Tel: +1 614 447 3661); for signup information fax +1 614 447 3713.

DATABASES

A series of computer disks containing composition and property data for a range of materials is being distributed through the UK Department of Industry's Materials Matter programme. Copper and Copper Alloys contains information about the compositions and properties of standard coppers and copper alloys based on data from the UK's Copper Development Association (CDA); a second disk covers copper-nickel alloys and aluminum bronzes, and includes data from CDA, Nickel Development Institute and Aluminium Bronze Advisory Service. The disks are produced by Engineering Information Co Ltd through the MATUS database system. CONTACT: CDA, Orchard House, Mutton Lane, Potters Bar, Herts EN6 3AP, UK; or Engineering Information Co Ltd, 15/17 Ingate Place, London SW8 3NS, UK.

STANDARDS

The name and scope of activities of ASTM Committee E49 have broadened to include chemical data, and work is well underway on the development of a standard chemical structure exchange protocol based on the SMD format. In the materials area, the following ASTM Standards related to the computerization of property data are in print or on the way: E1313-90 Guide for the Development of Standard Data Records for Computerization of Material Property Data; E1314-90 Practice for Structuring Terminological Records Relating to Computerized Test Reporting and Materials Designation Formats; E1338-90 Guidelines for the Identification of Metals and Alloys in Computerized Material Property Databases; E1339-90 Guidelines for the Identification of Aluminum Alloys and Parts in Computerized Material Property Databases; E1407-91 Guidelines for Materials Database Management. CONTACT: ASTM Staff Manager Teresa Cendrowska, 1916 Race Street, Philadelphia, PA 19103, USA, Tel: +1 215 299 5546.

DIRECTORIES

The second edition of the UK Directory of Materials Information Sources contains almost twice the number of entries as the first edition, and includes many materials suppliers and processors and a new section on standards. The new edition also covers sources of information from abroad that are readily available to users in the UK. CONTACT: The Design Council, 28 Haymarket, London SW1Y 4SU, UK.

CALENDAR

For a number of reasons, two events listed in the February issue of the Newsletter have had to be cancelled or postponed: Materials Data for Computer Aided Engineering, due to be held in Petten, The Netherlands, in March 1991; and Materials Database Experience—Opportunities for International Exchange, scheduled for April 1992 in Moscow, USSR.

CALENDAR (cont'd.)

The Third International Symposium on the Computerization of Materials Property Data will be held in Cambridge, UK, in September 1991. This important event will focus on Standards and Data Representation; Standards and Database Development; Expert Systems and Materials Databases; Data Issues for Engineering Materials; and Industrial Applications. The event will include demonstrations of a number of data systems.

9-11 September 1991, Cambridge, UK

Third International Symposium on COMPUTERIZATION OF MATERIALS PROPERTY DATA. CONTACT: Tom Barry, NPL, Teddington, Middlesex TW11 0LW, UK; or John Rumble, Jr., NIST, A323 Physics Building, Gaithersburg, MD 20899, USA.

Future meetings of ASTM Committee E49 have been scheduled as follows:

4-7 November 1991: San Diego, CA, USA 18-20 May 1992: Pittsburgh, PA, USA 16-18 November 1992: Miami, FL, USA

May 1993: Atlanta, GA, USA

November 1993: Fort Worth, TX, USA

EDITOR: W.G. Jackson. EDITORIAL OFFICE: The Institute of Metals, 1 Carlton House Terrace, London SW1Y 5DB, UK. Tel: +44 1 839 4071; Telefax: +44 1 839 2289; Telex: 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

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

COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY



CODATA, 51 bd. de Montmorency, 75016 Paris