INTERNATIONAL COUNCIL OF SCIENTIFIC UNIONS
COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY
ISSN 0538-6918

42 CODATA NEWSLETTER

NOVEMBER 1987

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.

CODATA Thermodynamic Tables

The "CODATA Thermodynamic Tables - Selections for Some Compounds of Calcium and Related Mixtures: A Prototype Set of Tables" have been sent to over 200 thermodynamicists around the world for advice and comment on behalf of the CODATA Task Group on Chemical Thermodynamic Tables which prepared it.

The book is the beginning of a CODATA series of tables of thermodynamic properties. Another book, the "CODATA Key Values for Thermodynamics," is at the printers and the Task Group is now working on a set of tables on compounds and two component systems of iron. Their goal is to cover the periodic table. However, at this stage, the Task Group would like the advice of the thermodynamic community. The tables represent several departures from previous tables both in terms of what is included and how they were produced. The differences are discussed in more detail in the first three chapters of the book but fall roughly into three categories: design, content, and modus operandi.

Comments are sought on the design of the tables for a particular set of compounds including the arrangement of values and their documentation, the contents of these tables, the type of material presented, the coverage of each topic, the satisfaction of the analysis, and the level of documentation.

Operationally, the tables are internationally sponsored and internationally produced. In spite of the decentralized mode of production, they are thermodynamically consistent. Details are given in Chapters 1 and 3. The methodology allows for different sections to be prepared simultaneously and for experts in specific areas to work cooperatively with others in producing the overall tables. Finally, the task group seeks reactions to this plan for producing a comprehensive set of tables, because they hope to speed up the work by increasing the number of participants.

Despite the unusual emphasis on criticism and advice, the Tables are no mean achievement. They comprise a very generous selection of calcium compounds and systems in a variety of states of aggregation.

Forty one tables of thermal functions (heat capacity, entropy, enthalpy, and Gibbs energy functions) for those compounds of magnesium, calcium and potassium for which the properties have been evaluated in this work are included. Twenty four tables of thermal functions are given for auxiliary substances. Each table spans the temperature range 0 to 4000 K, to the extent that data are available. Formation properties at 298.15 K (enthalpy and Gibbs energy of formation) are given for 68 compounds of calcium, magnesium and potassium plus the relevant values for 54 auxiliary substances. For each of the three systems the data given are mixing properties (enthalpies and Gibbs energies), partial molar or excess properties (activities, enthalpies, heat capacities) and phase transformation data. All selections of data are accompanied by uncertainties, and are documented. Detailed lists of measurements used in the evaluations are given for the thermochemical, alloy, and aqueous solution sections.

Nucleic Acid and Protein Sequence Data Workshop

The first CODATA Workshop on Nucleic Acid and Protein Sequence Data was organized by Dr. Rita R. Colwell (University of Maryland), Scientific Advisor to the White House.

The 130 participants, 15 of whom came from the EEC countries (5 from France), attended this meeting at the National Bureau of Standards, Gaithersburg, Maryland.

Dr. G. CAHILL, from the "Howard Hughes Medical Institute" opened the session, around the theme of mapping the human genome, a subject which was not directly treated again during the session.

The major part of the talks turned to present problems of the two main data banks on DNA sequencing, i.e., "GenBank"--Los Alamos National Laboratory and EMBL (Heidelberg, Federal Republic of Germany)--and on some other protein sequence databases. These problems are:

- explosion of the quantity of data to be entered
- impossibility to actually control the quality of data
- compatibility with other banks
- nomenclature
- relations with bibliographical data banks.

Solutions have been evoked, such as:

- increase of interfacing between source banks and users, thanks to networks and specified services (BIONET)
- decentralization of the sites for treatment of data
- participation of the biotechnology industry
- use of compact disks (CD-ROM)...

Only a few talks precisely stated the terms of this issue:

● C. BURKS, from Los Alamos National Lab, drew a parallel between the progress of sequencing techniques and the flow of data--7 million nucleotides per year at present time and 15 million in 1990. The human genome, with its 7 billion nucleotides should be deciphered about 2005.

● R. T. WALKER, B. CRAWFORD, and some other editors have clearly stated that they will no longer publish sequences simply for their description. They have also asked the two major data banks to permit reserved access to their data to enable their referees to control the sequences proposed by the authors; sequences that would have been initially registered in these banks.

● A. LESK (Laboratory of Molecular Biology, Cambridge, U.K.) presented his research on utilization of the primary data of these data banks, i.e. the three-dimensional structures of proteins and DNA.

● A. BUSSARD (CODATA) stressed the role of his organization as a privileged site for working on the basic problems, i.e., standardization of the data exchange formats and definition of the nomenclature principles. The sequence data banks will inevitably become basic information sources for biological research in the next few years and no longer sources secondary to scientific literature as is presently the case. This talk was designed to preclude the creation of an exclusively American ad hoc committee.

The National Library of Medicine coordinates dissemination of information on biotechnology in the United States, as was confirmed by Dr. G. Cahill in his welcoming address. At the time of the meeting, however, Congressional funding for this NLM program was yet to be approved, according to the Director, Dr. Lindbergh.

Connecting Computers to International Networks

The Canadian Fusion Program (Tokamak de Varennes, near Montreal, Canada) recently linked its local computer network (ETHERNET) to other computers through international network facilities. The connection is effected by linking the local ETHERNET network to the DATAPACK network of Bell Canada. The connection of DATAPACK to the other major networks around the world (TYMNET in U.S.A., TRANSPACK in France, DATEX in Germany, etc...) is obtained through the X-75 Protocol developed by CCITT (Comité Consultatif International Télégraphique et Téléphonique) in which Téléglobe Canada participated. The object of X-75 is to allow inter-networking; it provides a gateway for a user to communicate through multiple networks with another user. It allows data transfer on interntional circuit between packet switched data networks. Access to the huge scientific and database libraries of the fusion programs around the world have been made available to the Canadian Fusion Program. Execution of scientific programs on the computers of the MFE (Magnetic Fusion Energy) center at Livermore, California or on other computers around the world, and reception of the resultant graphic display are achieved with a high level of scientific quality according to Mogdi M. Shoucri of Tokamak de Varennes.

Interlinking of international networks for electronic mail, as well as scientific, educational, and research purposes is described tersely by J. M. Hudson in J. Molecular Graphics, Vol. 5, No. 1, pp. 57-58 (1987).

Special Chinese Geothermodynamics Group Founded

The Chinese Society of Mineralogy, Petrology, and Geochemistry has approved the founding of a Special Group on Geothermodynamics under their auspices to study relevant geothermodynamic matters and to interact with international endeavors in the critical evaluation of geothermodynamically relevant data and thus to provide within China a nucleus in this currently important area.

The Group numbers eight and their professional affiliations are:

Zhang Taok, Consultant, Secretary for Academic Affairs, Division of Earth Sciences, Chinese Academy of Sciences (and Chinese National Committee for CODATA), Beijing

Guo Qiti, Group Head, Associate Research Professor, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang

Lin Chuanxian, Research Professor, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang

Zhou Xunruo, Professor, Chinese University of Geology, Beijing

Zhang Ronghua, Associate Research Professor, Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing

Jiang Peimo, Associate Professor, Department of Geology, Peking University, Beijing

Yin Hui-an, Associate Professor, Graduate School, Chengdu College of Geology, Chengdu

Cheng Weiji, Associate Professor, Department of Earth and Space Sciences, University of Science and Technology of China, Hefei

MARCELLO CARAPEZZA

(1925 - 1987)

Professor Marcello Carapezza, Italian delegate to CODATA since 1974 and member of the CODATA Executive Committee, lost a two year struggle against cancer on September 1, 1987 at the age of 62. A long-time member of CODATA and the international scientific community, Professor Carapezza was a warm, caring geochemist whose life had been devoted to research in the hope of alleviating the human tragedy associated with natural disasters such as volcanic eruptions, earthquakes, and acid rain.

A native of Sicily, where science has only been an active word in the vocabulary for the last hundred years, Marcello Carapezza received his doctorate at the University of Palermo in 1959. From November 1959 to November 1969 he taught at the University of Bologna, interrupted in 1962-63 by a term as Associate Professor of Experimental Geochemistry at Pennsylvania State University. From 1969 he worked until his death at the University of Palermo as Professor of Applied Geochemistry. From 1970-72 he was President of the Faculty of Science and from 1982 to 1984 served as Vice-Rector of the University of Palermo.

At the time of his death he was the Director of the Institute of Geochemistry of Fluids of the National Research Council; President of the Advisory Committee for Earth Sciences of the Ministry of Public Education; Member of the Commission on Natural Hazards attached to the Ministry of Civil Protection; leader of the surveyance of active volcanoes for the national group on volcanology; Delegate from the Sicilian Region to the Public Society of the Strait of Messina in which he chaired the Committee on Geoseismotectonics; and member of the Regional Committee for Environmental Education.

In international activities he was a member of the Executive Committee of the following scientific unions: Geochemistry and Cosmochemistry Association (GCA), International Lithosphere Program (ILP), and the Committee on Data for Science and Technology (CODATA) where he was elected Chairman of the Task Group on Monitoring Active Volcanoes during its international meeting in Israel in 1984.

In 1982 the President of the Republic, Sandro Pertini, awarded him the Gold Medal for Merit in Culture, Science and Art. In 1984 he received the Mauroleco Prize in Messina.

A consultant to the CEC on problems of the environment, his scientific activities led to publication of over 100 articles (several of which were translated into various languages), from the compilation of geological maps in the area of Sardinia and Sicily to a large number of articles in newspapers such as The Corriere dellaSera, Il Gioronale di Sicilia, L'Ora, Panorame, L'Europea, etc. and the American New York Times. In addition he was involved in the publication of scientific books and wrote the prefaces to several others.

He organized international scientific conferences, directed the International School of Earth Sciences of the Majorana Center in Erice, presented scientific papers in U.S.A., U.S.S.R., Great Britain, France, Federal Republic of Germany, Israel, Japan, and Venezuela.

He set up bilateral agreements for scientific cooperation with the U. S. Geological Survey, Reston, Virginia, U.S.A.; the Academy of Sciences of the U.S.S.R., Moscow; the Lomonosov University, Moscow; and the Institute for Chemistry of Earthquakes, Tokyo, Japan.

A member of several Italian and foreign academies, he received scientific prizes from Italy, France, and the U.S.A., and was made an honorary citizen of Gibellina (Trapani).

His research included studies of large natural processes such as volcanoes, contributions to the study of the lunar field, studies related to atmospheric degradation of historic monuments, prospecting for geothermic resources, and above all, methods of studying geochemical activities of volcanoes including their continuous monitoring.

In 1987, the "Year of the Environment," the Minister Granelli awarded him in Milan a scientific prize for his important contribution to the protection of the environment and particularly recognized his major contribution to the forecasting and prevention of earthquakes and his studies relating to volcanoes and atmospheric pollution.

CODATA extends heartfelt condolences to his wife, Ginevra, and his four children, Attilio, Fabio, Marco, and Marisa.

--Phyllis Glaeser

IUIS not IUBS!

CODATA Newsletter No. 41, page 6 correctly identified the new delegate from the International Union of Immunological Societies, but the heading should have reflected the text and read IUIS (not IUBS).

Fundamental Constants--Cohen

To the "Comments" of Dr. M. Grabe, CODATA Newsletter 41, p. 6, July 1987, Dr. E. Richard Cohen, former Chairman of the involved CODATA Task Group replies:

Dr. M. Grabe
Braunschweig, F.R.G.

September 22, 1987

Dear Dr. Grabe:

I have carefully read your paper, Principles of "Metrological Statistics," published in Metrologia and your comments on the 1986 adjustment of the Fundamental Physical Constants in the CODATA Newsletter. Although there are significant differences between the treatment that you propose and the treatment utilized by Dr. Taylor and me, I do not think that we are in disagreement about basic statistical procedures.

Your treatment is directed towards giving a confidence interval for the true value of an observed (or inferred) variable. My viewpoint is that a confidence interval is an extremely poorly defined quantity. Because the tails (extremes) of a distribution function are themselves not well defined, it becomes increasingly poorly defined as the confidence level increases.

Although the discussion in CODATA Bulletin 63 did not specifically talk about confidence intervals, I am in essential agreement with you that confidence intervals should be calculated from formulas such as equation (2.6) of your Metrologia article; i.e.,

$$\hat{x} - t_\nu \hat{\sigma}_x < x < \hat{x} + t_\nu \hat{\sigma}_x \qquad (1)$$

where $\hat{\sigma}_x$ is the estimate of the standard deviation of the distribution function associated with the variable x, and t_ν is the Student parameter associated with the specified confidence level. Note here that I am using x with the same meaning that you use the symbol $\bar{x}$. We calculated the effective degrees of freedom in agreement with B. L. Welch,

$$\sigma^2 = \sum \sigma_i^2, \qquad \sigma^4/\nu = \sum(\sigma_i^4/\nu_i). \qquad (2a,\ 2b)$$

Equation (2) may be considered to be a recursive definition; the components σ_i^2 of σ^2 and their associated degrees of freedom ν_i are themselves to be calculated from similar equations, if necessary. The primary quantities required for the evaluation of confidence intervals are estimates of the mean and the variance, and the number of degrees of freedom associated with the variance estimate. If these cannot be stated, a confidence interval cannot be given. The Student distribution is based on the assumption of a normal distribution for the underlying stochastic data. If this is not the case, the intended confidence interval is not correctly given by equation (1). Certainly, no precise confidence level can be given for the $(x - u_x, x + u_x)$ interval with $u_x = t_\nu \sigma + f_s$ from equation (2.7) of your paper. One cannot quote a confidence interval if the statistical estimate is simply increased by a "safe" estimate of the magnitude of unevaluated error.

We have followed the BIPM procedure recommendation to treat 'systematic' error in the same manner as 'random' error (and have used essentially the same concepts for years prior to the BIPM proposal) because there is no valid alternative for the general purposes with which we are dealing. The BIPM recommendation, if I may give it my own interpretation, does not distinguish between 'random' and 'systematic' errors because such a distinction is in fact impossible.

'Random' errors are recognized as errors that are unpredictable and uncorrelated from measurement to measurement. 'Systematic' errors are undetected in the measurement process because they are repeated from trial to trial with a correlation coefficient of 100%. From the analytic standpoint, however, error is described by a (possibly unknown) distribution function. For a finite sequence of measurements, this can be decomposed into a fixed component (non-zero mean) with zero dispersion (zero variance) and a random component with zero mean and finite variance. In practice, however, this sharp dichotomy does not exist; there is no unique decomposition of the error into components, and the errors of two different trials are correlated with an arbitrary correlation coefficient that may vary for every pair of trials.

Experimental error is usually described in terms of what the experimenter perceives to be the sources or the origins of the errors. It is impossible to observe the constant component of the so-called "systematic' errors. However, if a 'systematic' source has a variable component it will contribute to the observed dispersion of the measurement results. Any decomposition of errors on the basis of source or origin must consider the extent of the correlation (from -1.0 to +1.0) among the components.

When the uncertainties of the systematic errors are expressed in terms of intervals it is still necessary to consider the question of how to combine components, including possible correlations. Only if two errors have a correlation coefficient of 1 may the two error limits be added. (Two 'random' variables will also combine so that the uncertainty of the sum is the sum of the uncertainties when the correlation coefficient is 1.) Furthermore, the assumed 'limit' interval is uncertain, and the fact that the limit intervals of two measurements of the same quantity do no overlap indicates that the limits have been incorrectly evaluated. Thus, even if one were to treat 'random' and 'systematic' uncertainties differently, it would be necessary to include a statistical treatment of the magnitudes of the 'limits of error.' It seems to me that this leads to even murkier realms than the uncertainties associated with the BIPM recommendations.

Fortunately, the additive properties of the characteristic function and the Central Limit theorem come to our aid here: the second and fourth moments of the distribution function (or if one wishes, the variance and the degrees of freedom) of a stochastic variable can be estimated in a straightforward manner from equations (2a) and (2b). Since these equations embody the only statistically valid statements that can be made, I believe that we are correct in quoting only these quantities.

My viewpoint is that the precise distribution of errors is unknown, and we therefore avoid the issue by dealing only with the quantities that are available to us. Our analysis is intended to provide the mean value, the variance, and the effective degrees of freedom (equivalent to giving the fourth moment) of a stochastic variable or of a set of data. The user can then construct whatever confidence intervals may be deemed appropriate using whatever further statistical processes are desired. Questions of probabilities and of confidence intervals are important, but these are the next (and independent) step, and involve their own assumptions and uncertainties. We do not consider it appropriate to extend our analysis to cover these questions because, to a large extent, such questions are problem-dependent and outside the range of our objectives.

I would also like to point out that it is incorrect to say, as you do in your note in the CODATA Newsletter, that "the term 'error' has disappeared from official linguistic usage". This is by no means true; rather the attempt is being made to distinguish two related but very different concepts. Physical measurements are subject to 'error"; one never knows what the errors in a measurement are, otherwise they would, of course, be corrected before giving the final result. The term 'uncertainty' is used to denote an estimate of the magnitude of the 'error'; the BIPM recommendation is to use the standard deviation of the observed or assume distribution of errors as a measure of this uncertainty.

Sincerely yours,

--E. Richard Cohen
Distinguished Fellow
Rockwell International Science Center

CODATA Books

Banques De Données Factuelles Sur Les Materiaux, (Actes de la journée d'étude) edited by Bernard Marx. (a)

Books and Computer Services

IUPAC Handbook 1985-87. (b,c)

Compendium of Chemical Terminology: IUPAC Recommendations, by V. Gold, K.L. Loening, A.D. McNaught, and Pamil Sehmi. (b,d)

Quantities, Units and Symbols in Physical Chemistry, by I. Mills, T. Cvitas, K.H. Homann, Nicola Kallay, and K. Kuchitsu. (b,e)

Recommended Reference Materials for the Realization of Physicochemical Properties, by K.N. Marsh. (b,f)

International Thermodynamic Tables of the Fluid State, Volume 9: Oxygen, by W. Wagner and K. Marjorie De Reuck. (b,g)

The Chemical Thermodynamics of Organic Compounds by Daniel R. Stull, Edgar F. Westrum, Jr. and Gerard C. Sinke. (h)

CODATA Calendar

1988

March

3- 5	CODATA Task Group on Coordination of Protein Sequence Data Banks, Paris, France
3- 5	32nd CODATA Executive Committee, Paris, France

July

21-	CODATA Task Group on Geothermodynamic Data, Strasbourg, France
25-28	First International Symposium on Thermodynamics of Natural Processes, Strasbourg, France

September

26-29	11th International CODATA Conference, Karlsruhe, F.R.G.
30-Oct.1	CODATA General Assembly, Karlsruhe, F.R.G.

Footnotes

(a) CODATA, Paris, France, 1987. ISBN 2-906828-00-6. Price: 175 Fr. Francs. Available from CODATA Secretariat in Paris France.

(b) Blackwell Scientific Publications, Inc., P.O. Box 50009, Palo Alto, CA 94303-9952, or Blackwell Scientific Publications, Ltd., Osney Mead, Oxford, U.K., OX2 OEL.

(c) This Handbook, produced biennially as a supplement to "Chemistry International" provides detailed information on IUPAC, its personnel and IUPAC bodies. Significantly larger than previous Handbooks, its contents include IUPAC Officers and Secretariat; membership of all IUPAC Committees with full alphabetical index of names and addresses; listing of all official current programmes of the various committees; and detailed information on the IUPAC Affiliate Membership Scheme. 1986. 262 pages. US $32.00. ISBN 0 269 8757.

(d) This Compendium is an alphabetical listing of terms used in chemistry, with definitions recommended by various IUPAC commissions dealing with nomenclature and terminology published through the end of 1985. The definitions are truly authoritative, being recommendations agreed upon by committees of experts after consultation with the chemical community. It is intended to be ongoing as some important areas of chemistry have not yet been the subject of IUPAC recommendations, and thus, the first step towards a comprehensive and authoritative chemical dictionary. February, 1987. 464 pages. Cloth: 45 pounds/US$69.60, ISBN 0 632 01765 1. Paper: 29.50 pounds/US$48.45, ISBN 0 632 01767 8.

(e) The first edition of the **Manual of Symbols and Terminology for Physicochemical Quantities and Units** (the 'Green Book') of which this is the direct successor, was prepared for publication on behalf of the Physical Chemistry Division of IUPAC in 1969. This book had as its objective the clarity, precision, and wider agreement in the use of symbols by chemists in different countries, among physicists, chemists and engineers, and by editors of scientific journals. The second and third editions of the 'Green Book' were revisions to take account of various developments in nomenclature. This present volume is a substantially revised and extended version of the earlier editions, with a slightly simplified title. These extensions include some of the material previously published in appendices, all the newer resolutions and recommendations of the International Union of Pure and Applied Physics (IUPAP) of 1978 and the International Organization for Standardization. The table of physical quantities has been extended to include defining equations and SI units for each quantity, and the style has been changed from being a book of rules towards being a manual of advice and assistance for the day to day use of practicing scientists. Summer, 1987. About 126 pages, illustrated. Paper about 9.80 pounds/US$15.00, ISBN 0 632 01773 2.

(f) This definitive volume, containing contributions from internationally acclaimed experts, provides a comprehensive review of the various materials which can be used as references for checking the correct operation or calibration of equipment used for physicochemical measurement. The important areas covered include density, viscosity, enthalpy, thermal and electrolytic conductivity, optical rotation and refraction, and relative molar mass. It will become an essential reference text for all physical and analytical chemists and chemical engineers, and those working in standards laboratories. Summer, 1987. About 470 pages, 6 illustrations. About 56 pounds/US$70.00, ISBN 0 632 01718 X

(g) The main part of this book consists of tables of the equilibrium thermodynamic properties of oxygen for both the gas and liquid phases, and along the saturation and melting curves. The limits of the tables are 54 k to 300 k for pressures up to 90 MPa. The properties listed are volumes, entropy, enthalpy, isobaric heat capacity, compression factor, fugacity-pressure ratio, Joule-Thomson coefficient, ratio of heat capacities and speed of sound as functions of pressure and temperature, and pressure, entropy, internal energy, and isochoric heat capacity as functions of density and temperature. The text gives a critical assessment of the accuracy of all the available experimental data and gives comparisons of the tables with the experimental data. The equations used to prepare the tables are given in a form suitable for computer programming. Autumn, 1987. About 256 pages, 27 illustrations. Cloth about 30 pounds/US$37.50, ISBN 0 632 01476 8.

(h)Robert E. Krieger Publishing Company, Krieger Drive, Malabar, FL 32950. Original edition, 1969; reprint with corrections, 1987. About 865 pages. ISBN 0-89874-706-6.

Letter to the Editor

Sir,

You very kindly published a letter by me (Newsletter No 34, October, 1985) commenting on the Executive Secretary's enjoyable article about the Hotel de Noailles and Auteuil. One line of my original typescript was inadvertently omitted from the printed version and the error was not spotted at the proofreading because, in spite of the omission, the sentence, grammatically, made sense.

I give the correct version with the omitted passage in *italics*: 'He (Count Rumford) spent most of his life partly in Munich where he created the "English Garden," established workhouses to make the numerous beggars work, feeding them on "Rumfordsche Suppe" (1000 calories per day) and where he became a count of the Holy Roman Empire, and *partly in London where he was Knighted, elected a Fellow of the Royal Society and* founded the Royal Institution of Great Britain.'

I should be grateful if you could publish this note and thus ensure that your readers realize that, unlike the "English Garden," the "Royal Institution" is not in Munich but in London.

Yours faithfully,

--N. Kurti

Factual Databanks

A small, but interesting volume entitled "Banques de Donnees Factuelles sur les Materiaux (Actes de la Journee d'Etude)" summarizing the action at France's Study Day--November 20, 1986--under the auspices of CODATA FRANCE DBMIST/MRES has appeared recently.

As its title implies, it is concerned with material data base development incorporating almost two score presentations of experts from several countries on both sides of the Atlantic. Moreover, annexes cover such useful topics as a Code of Practice (EEC's MDB Demonstrator), a sampling of the forthcoming materials section of the Inventory of Data Sources in Science and Technology, a list of actual data banks on materials (including those cited at the Study Day) and the (more than 200) participants.

The 184 page volume is edited by Bernard Marx of DBMIST and bears the mark of the Comite des Donees pour la Science et la Technologie under the presidency of Dr. H Viellard. More details on its procurement will be found on page 5.

CODATA Conference Proceedings

Another new -- and better -- summary of the broad spectrum of data handling and dissemination has appeared in compact form in the new volume "Computer Handling and Dissemination of Data" which comprises the proceedings of the Tenth International CODATA Conference at Ottawa, Canada, 14-17 July 1986.

The editing has been done under the direction of CODATA's Executive Secretary, Phylis S. Glaeser and the publication by North-Holland.

A CODATA Bulletin of the same title but dealing with otherwise selected papers, CODATA Bulletin 64 (December, 1986), is listed in CODATA Newsletter No. 40 and should not be confused with the present tome which is listed among CODATA publications on page 7 of CODATA Newsletter 41.

The 82 papers contained in this volume have been selected from more than 180 papers presented at the 10th International CODATA Conference. These papers cover some of the general aspects of computer applications in the management, dissemination, and use of factual scientific data, as well as the status of database design, manipulation, and development in materials science, chemistry, physics, astronomy, geoscience, environmental science, molecular biology, biotechnology, nutrition and agriculture disciplines.

The Conference brought together 300 experts in the evaluation, dissemination, and management of numerical databases in all fields of science and technology. The central theme of the Conference, "Computer Handling and Dissemination of Data," was brought out in sessions devoted to data validation by statistical techniques, general approaches to database design, and options for computer-based dissemination of data to the scientific community. A strong theme throughout the Conference was the growing database management capability of microcomputers and the prospect for the dissemination of databases in PC compatible formats for use by individual scientists.

Thermodynamics of Alloys

A new volume, "Statistical Thermodynamics of Alloys" by Nev Gokcen covers solutions of metals in metals and both the statistical thermodynamic aspects and numerical data values of metalloids in metals in depth for discussion of the procurement of thermodynamic values from phase diagrams, as well as the quantum statistical mechanical aspects (often postulative as well as inductive and deductive).

Solution theory supplemented by treatment of order-disorder phenomena are provided together with values for the actual numbers of configurations in highly ordered structures. The extensive treatment of interstitial solutions--particularly for hydrogen storage in metal and alloys--is an important feature of the text. Other renewable energy relevant material concerns the conversion of solar energy into hydrogen and/or electrical energy.

The extensive appendices contain recent and evolving theories of alloy phase formation, methods of estimation and correlation of thermodynamic properties (including those in dilute solutions, the Engel-Brewer approach in Appendix A, and those concerning the estimations of the enthalpy of alloy formation by Miedema's group). Extensive tables--and phase diagrams--are included for the correlations of excess partial molal enthalpies and entropies in solution. Hence the development and results are indeed of interest not only to metallurgists, material scientists, and ceramicists, but to geoscientists as well. The author is to be commended on what must have been a large (long) and arduous achievement.

Further details concerning the production and procurement of the book are given on page 7, CODATA Newsletter No. 41.

Multisatellite Mapping

The Task Group meeting in Beijing and Xi'an (China) 1-14 September 1987 reviewed its organization and plans for 1987-1988. The main participants were Dr. C. Bardinet, Chairman (CTAMN, France), Prof. G. Gabert (BGR, FRG), Prof. Chen Yiyun (RSGS-Beijing), and Prof. Wan Zheng-Ming (IRSA, Academia Sinica, Beijing). The organization of the Task Group has been reviewed relative to the objectives of CODATA China as explained by Prof. Mooson Kwauk and Ms. Hu Yaruo, Secretary.

The CODATA Chinese National Committee plans to establish a Multisatellite Thematic Mapping Task Group in China (MTMTGC) in order to implement research cooperation with the CODATA Task Group (MTMTG) during 1987-1988. The Task Group will consist of four Chinese members and four members of the CODATA Task Group.

The Task Group plans a presentation at the 1988 Karlsruhe Conference on Multisatellite Imagery Applications in China. The lecture will consist of a comprehensive presentation on Chinese Remote Sensing Projects using Multisatellite Imagery (prepared by the Chinese members), and joint research involving integration of high resolution imagery including TM and SPOT classifications.

The research areas selected include: (1) Tangshan city in Hebei Province and Nanbi county, (2) the Loess plateau of Shaanxi Province, (3) Northern Xianjiang Province. In Hebei and Shaanxi the research will concern mainly natural resources classification (TM and Spot) and mapping (geomorphology and land resources); whereas in Xinjiang, it will consist mainly of comparative geological studies based on NOAA and MSS imagery.

In Beijing, discussions concerning future work took place at: 1) the IRSA (Institute of Remote Sensing Application, Ac. Sin.) with Prof. Yang Shi-Ren, Prof. Wan Zheng-Ming and Prof. Chen Yiyun; 2) the TM Remote Sensing Ground Station with Prof. Lee Zhi-rong and Prof. Chen Yiyun; 3) the NOAA Ground Station with Dr. Liu Chen, Dr. Sun Zi-Yu and Prof. Wan Zheng-Ming.

In Xi'an similar discussions took place at the Institutes of Geography and Geology (Northwestern University) with Dr. Yuan Kansheng and Dr. Dai Mo, at the Photogrammetry and Remote Sensing Center of the Ministry of Coal Industry with Dr. Liu Yonglian, Mr. Song-Li, Mr. Yao Ding-Yi and Mr. Gao Yi. C. Bardinet and G. Gabert met with Prof. Wang Yong-Yan, a world-renowned loess soil specialist.

Applications of remote sensing to land use and land resources mapping such as forest fire surveying using NOAA, detection of geomorphological features using TM and SPOT, detection of geological structures affecting coal mining, and geographical information system using TM, MSS and SPOT were discussed.

-- Claude Bardinet

Microbiological Data Coding

Morrison Rogosa and Micah I. Krichevsky of the National Institutes of Dental Research of the National Institutes of Health, and Rita R. Colwell of the University of Maryland have assembled an open ended system in which computer techniques facilitate encoding, entry, management, and analysis of microbiological data derived from the study of bacteria, algae, fungi, and protozoa. Many others have contributed expertise and productive assistance, and the 310 page volume has now been realized under CODATA sponsorship as a volume in the Springer Series in Microbiology.

Further details concerning its availability may be found on page 7 in CODATA Newsletter # 41.

TASK GROUP ON MATERIALS DATABASE MANAGEMENT

MATERIALS DATABASE NEWSLETTER October 1987, Number 1

The purpose of this Newsletter is to provide an overview of national and international developments relating to all aspects of materials property databases, and to promote cooperation with respect to standards for database design, management and use. The Editor welcomes contributions and comments from database builders, operators and users. There are no restrictions on the reproduction and distribution of the contents of this Newsletter.

NETWORKS

The **CEC's DEMONSTRATOR PROGRAMME ON MATERIALS DATABANKS** is the first step in the European Commission's aim of promoting public availability of materials property databanks. The Demonstrator Programme, which is designed to become a pilot to a future European system, will integrate most of the European online databanks in a way which is expected to provide easier access, better transparency, more comprehensive services and a better response to user needs. The programme will involve harmonization of participating databanks and the development of a Multilingual User Guidance System. Seminars will commence late in 1987 and Demonstrator Workshops early in 1988. SOURCE: Echo News, 1987, No. 1; FURTHER INFORMATION: Mr. G. Steven, Commission of the European Communities, DGXIII/B, Bat. J. Monnet, Plateau du Kirchberg, L-2920, Luxembourg.

The **NATIONAL MATERIALS PROPERTY DATA NETWORK** is an autonomous corporation, set up in the U.S. but international in scope, which will provide access to a wide range of data sources through a common gateway. A major pilot programme was established in 1986 with the aid of a number of government and industrial sponsors. Successful online demonstrations of the pilot network were carried out for sponsors in May this year. Three sources are included in the pilot network: portions of Military Handbook-5 (MIL-HDBK-5) and the Aerospace Structural Materials Handbook; and STEELTUF, a large bank of toughness test results for about 50 steels. The MPD Network has also concluded an agreement with STN International to have the network carried on that system, though no schedule has been announced. SOURCE: MPD Network Newsletter, April 1987. FURTHER INFORMATION: J.G. Kaufman, MPD Network, Inc., 2540 Olentangy River Road, P.O. Box 02224, Columbus, OH, U.S.A.

STANDARDS

ASTM COMMITTEE E49 on Computerization of Materials Property Data was established in 1986 to promote knowledge and to develop standard classifications, guides, practice and terminology for building and accessing material property databases. Four sub-committees have been formed to cover the areas of Identification of Materials; Reporting of Material Property Data; Terminology; and Database Interfaces and Functionalities. SOURCE: ASTM Standardization News, Mar. 1987, p. 38.

VAMAS (VERSAILLES PROJECT ON ADVANCED MATERIALS AND STANDARDS) has established a Technical Working Area on materials databanks with the objective of identifying standardization activity related to the computerization of materials data. The report of the Working party will shortly be published, and will cover: The Need for Standards; Data Collection; Database Building, including Data Analysis; Access to Data; Computer Integrated Engineering and Expert Systems; and Organizations Active in Materials Data Standards. SOURCE: VAMAS Bulletin, 1987, (5), p. 11. FURTHER INFORMATION: Dr. J. Rumble, NBS, A323 Physics Building, Gaithersburg, MD 20899, U.S.A OR H. Kroeckel, CEC Joint Research Centre, NL-1755 ZG, Petten, The Netherlands.

In connection with the CEC's Demonstrator Programme, a **CODE OF PRACTICE** has been developed to provide guidelines for the functional operation of software and standards for uniformity and user friendliness. The specifications cover four main areas: Access Level, Host Operations Level, Systems Operation Level, and Contents Level. Copies of the Code of Practice are available from DGXIII/B in Luxembourg.

DIRECTORIES

The Institute of Mechanical Engineers has recently published a U.K. directory of **MATERIALS DATA SOURCES**, which includes a list of databases and materials selectors. The directory is available from: MEP Publications Ltd., PO Box 28, Northgate Ave., Bury St. Edmunds, Suffolk, U.K. IP32 6BW PRICE: 10.00 pounds.

DATABASES

The Institute of Metals and Materials Australasia have announced the availability of a computerized materials selection system, developed by Dan Phelan of Chisholm Institute of Technology. Initially the system will include data for Al and Cu alloys; steels, cast irons, and all common alloys will be added by the end of 1987. SOURCE: Materials Australasia, Mar. 1987, p. 23.

CALENDAR

29 Sept - 1 Oct 1987: Trier, Federal Republic of Germany

International Congress on **TERMINOLOGY AND KNOWLEDGE ENGINEERING.** CONTACT: C. Galinski, Ges. fur Terminologie und Wissenstransfer e.V., D-5500 Trier, Postfach 3825, Federal Republic of Germany.

2-3 Nov 1987: Philadelphia, PA, U.S.A.

1st International Symposium on **COMPUTERIZATION AND NETWORKING OF MATERIALS PROPERTY DATABASES.** CONTACT: Dr. J. Rumble, NBS, A323 Physics Building, Gaithersburg, MD 20899, U.S.A.

11-13 Nov 1987: Yokohama, Japan

ICIK 87, International Conference on **INFORMATION AND KNOWLEDGE.** CONTACT: Prof. S. Fujiwara, Kanagawa University, 3-27-1 Rokkakubashi, Kanagawa-ku, Yokohama 221, Japan.

26-29 Sept 1988: Karlsruhe, Federal Republic of Germany

11th International CODATA Conference **SCIENTIFIC AND TECHNICAL DATA IN A NEW ERA.** CONTACT: DECHEMA, Abt. Tagungen, POB 97 10 46, Theodor-Heuss-Allee 25, D-6000, Frankfurt am Main, Federal Republic of Germany.

EDITOR: W. G. Jackson. EDITORIAL OFFICE: The Institute of Metals, 1 Carlton House Terrace, London, U.K., SW1Y 5DB. Tel: +44 1 839 4071. Telex: 8814813. Telefax: +44 1 839 2289.

(Note also: Review of "Banques De Donnees Factuelles Sur Les Materiaux" on page 5.)

Editor pro tempore: Edgar F. Westrum, Jr.
Department of Chemistry, University of Michigan,
Ann Arbor, MI 48109
Telephone: (313) 764-7357/ Telex: 8102236056

Associate Editor: Phyllis Glaeser
CODATA Secretariat, 51 Blvd. de Montmorency,
75016 Paris, France
Telephone: 4525-0496 / Cables: ICSU Paris 016/ Telex: 630553

Published four times per year (January, April, July, and October)

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