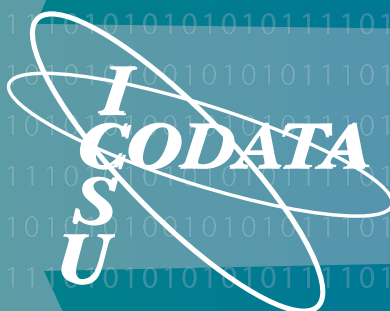
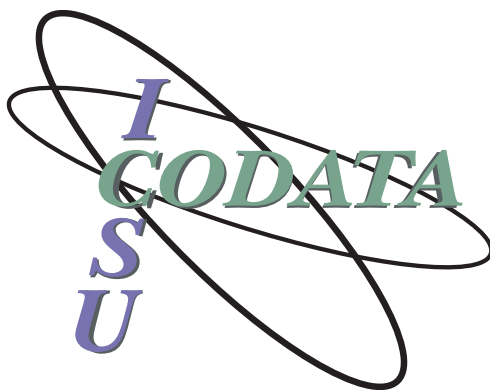


CODATA @ 45 YEARS

The Story of the ICSU Committee on Data for Science and Technology
(CODATA) From 1966 to 2010

David R. Lide • Gordon H. Wood





CODATA @ 45 Years

**The Story of the ICSU Committee on Data for Science and Technology
(CODATA)**

From 1966 to 2010

David R. Lide and Gordon H. Wood

Published by CODATA

© 2012 by David R. Lide and Gordon H. Wood
Some rights reserved



This work is licensed under the Creative Commons Attribution 3.0 Unported License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>.

ISBN: 978-0-9917424-0-0

Published by
CODATA
5 rue Auguste Vacquerie
75016 Paris, France
<http://www.codata.org/>

Layout by
El Designo Inc.
<http://www.eldesigno.ca>

Printed in Taipei

Table of Contents

Preface and Acknowledgments

Preamble	1
The 1966-1970 Period	5
The 1971-1980 Decade	11
The 1981-1990 Decade	19
The 1991-2000 Decade	29
The 2001-2010 Decade	41
CODATA in Retrospect	57
Appendix A – Sites of CODATA General Assemblies and Conferences	59
Appendix B – National Members – Years of Membership in CODATA	60
Appendix C – Scientific Union & Co-opted Members – Years of Membership in CODATA	62
Appendix D – Task Groups and Commissions – Years of Service to CODATA	64
Appendix E – CODATA Officers	69
Appendix F – Attendance at First International Conference - 1968	70

Preface

It is our pleasure to offer this account of CODATA's history to the scientific data community. While 45 years is not usually considered a benchmark for categorizing historical events, we thought it prudent to act now while the two of us, whose combined experience with CODATA spans this period, are available and able to write from first-hand experience. A further advantage was that several of the participants in important CODATA projects were able to provide input to our account. During the last ten years a couple of oral sessions celebrating accomplishments have been convened at CODATA meetings, but these were understandably incomplete and unpublished. Thus it was felt that a written record would serve not only to document past endeavours but to serve as a point of reference and inspiration for the future.

With the exception of the first five years, we have chosen to order the account by decades beginning in 1971, 1981, etc., with an internal parallelism in each decade to facilitate comparison. The primary written sources used were minutes of General Assemblies and Executive Committee meetings along with *CODATA Newsletters*, *Bulletins*, and other publications. As with any historical account, arbitrary decisions had to be made with respect to details concerning such items as events, people, accomplishments, and impacts. If anyone or any group has been offended, we apologize in advance.

As the reader will note, what CODATA has accomplished in these 45 years greatly exceeds the output that might be expected from its budget and paid staff. Thus, it is with pleasure that we dedicate this book to the hundreds of people who have volunteered their time and energy to participate in CODATA projects.

Acknowledgements

Clearly, compiling a story such as this is possible only with assistance and input from many others. In particular we wish to thank the following for their contributions.

Kathleen Cass and Cécile Carbonelle helped greatly by retrieving timeworn Newsletters and General Assembly Minutes from the archives and coordinating the publishing process.

Helpful reviews and comments were given by Lois Blaine, Robert Chen, Phyllis Glaeser, Micah Krichevsky, Steve Rossouw, John Rumble, Barry Taylor, David Watson, and Ronald Wigington.

Last, but not least, Linda Wood produced Appendix F from old files of questionable legibility and executed meticulous proof reading. Nonetheless, responsibility for any remaining errors remains with the authors.

David R. Lide
13901 Riding Loop Dr.
North Potomac, MD 20878

Gordon H. Wood
1800 Woodhaven Heights
Ottawa, Ontario K1E 2W2

Preamble

In the five centuries since modern science began, the quantitative measurement of natural phenomena, materials, and processes has played a central part in the evolution of scientific understanding. The data from Tycho Brahe's astronomical observations led to Kepler's solar system model; Galileo's measurements of pendulum swings were crucial to the development of Newtonian mechanics; data on the physical properties of chemical compounds were essential to our understanding of molecular structure and chemical reactions. Science could not have advanced to its present state if data from each generation had not been preserved and made available to subsequent generations.

As scientific research expanded, formal mechanisms for preserving data evolved, starting with the archival journal and expanding later into compilations of numerical data on specific topics. In the physical sciences comprehensive handbooks of data first appeared during the explosion of research in the second half of the 19th century. The *Beilstein Handbook of Organic Chemistry*, started in 1881, and the *Gmelin Handbook of Inorganic Chemistry* (1887) were – and remain today – basic resources for every chemist. The *Landolt-Börnstein Tables* (1883) covered physics, astronomy, geophysics, and physical chemistry. In geology, oceanography, astronomy, and meteorology various institutional repositories were established to preserve the relevant data.

Following the end of World War II, there was a rapid increase throughout the developed world in the number of scientists carrying out research. At the same time new, automated instrumentation made it possible to perform physical measurements much more efficiently. These two factors led to an exponential expansion in the amount of data published in the scientific literature and compiled in handbooks and repositories. By the early 1960s a number of scientific leaders began to realize that this deluge of data was swamping the traditional publication and retrieval mechanisms, and that there was a danger that much of it would be lost to future generations. When several of these leaders got together and agreed that an organized international effort was needed to improve the management and preservation of scientific data and to

facilitate coordination among interested groups throughout the world, the creation of CODATA was the outcome.

It is fair to say that the prime mover in this process was Frederick Rossini of the United States. He was a distinguished thermodynamicist who made his mark through very precise measurements of the energy changes in chemical reactions. He was a member of the National Academy of Sciences and, in the 1960s, was Vice President of Notre Dame University. Rossini obtained important support from Harrison Brown, Foreign Secretary of the US National Academy of Sciences and later President of the International Council of Scientific Unions (ICSU), and Allen Astin, Director of the National Bureau of Standards in the US.

Rossini discussed these data issues with colleagues in other countries and found a nucleus of scientists willing to take action. This group included:

- Boris Vodar of France,
- Sir Gordon Sutherland of the United Kingdom,
- Wilhelm Klemm of West Germany,
- Masao Kotani of Japan, and
- Mikhail Styrikovich of the Soviet Union.

They proposed a new organization under the umbrella of ICSU, and in June 1964 ICSU agreed to set up a working group, chaired by Harrison Brown, to study the proposal. The Working Group reported favorably and drafted a constitution, and in January 1966 the ICSU General Assembly in Bombay approved the creation of CODATA. The first governing body, which was called at that time the Bureau, consisted of the six people named above, with Rossini as President, Vodar and Sutherland as Vice Presidents, Klemm as Secretary-Treasurer, and Kotani and Styrikovich as Bureau Members.

Some words about the six Founding Fathers are in order. Frederick Rossini was a highly respected senior scientist. In addition to his research accomplishments, he had initiated a program at the National Bureau of Standards in the 1930s to compile and evaluate thermochemical data. He was a rather formal person, who mapped out a



Frederick Rossini

their data output to the CODATA Central Office for dissemination. Viewed in retrospect, this was not a realistic model, but his determination was instrumental in launching CODATA as an effective organization.



Boris Vodar

Boris Vodar was a Russian émigré to France whose special field was high pressure physics, but who had very broad scientific interests. He founded the Laboratory of High Pressures at Bellevue, France, and published over 300 papers on the properties of materials at high pressures and temperatures. In the process, he won many scientific honours. Beyond that, he was an expert on abstract paintings and a connoisseur of wine, planting his own vineyards in Burgundy. Those who attended the 1972 CODATA Conference in Le Creusot, which he organized, learned as much about wine as about data. He was passionate in his support of CODATA, and as second President did much to expand the membership.



Sir Gordon Sutherland

Sir Gordon Sutherland was a Scotsman who did early research in infrared spectroscopy and the application of quantum mechanics to the determination of the three-dimensional structure of molecules. He was Director of the National Physical Laboratory, the premier British Government physical science laboratory, from 1956 to 1964 and later became Master of Emmanuel College at Cambridge University. Sutherland was a top advisor to the British Government on science and education policy. Outside of science, he was a Trustee of the National Gallery and a connoisseur of Chinese art.



Wilhelm Klemm

Wilhelm Klemm was a distinguished German chemist, who made many contributions to inorganic chemistry during his career. He won numerous honours for his research from Germany, France, and other countries. He grew up in the tradition of the great German data compilations, such as the *Beilstein Handbook of Organic Chemistry* and *Gmelin Handbook of Inorganic Chemistry*, and was himself a contributor to the *Landolt-Börnstein Tables*.



Masao Kotani

Masao Kotani, a Professor at the Science University of Tokyo, began his scientific career in the early days of atomic physics and later moved into biophysics. At the time of CODATA's founding he was one of the senior scientific statesmen in Japan. Through his influence a number of other Japanese scientists became active in CODATA. He established the Task Group on Accessibility and Dissemination of Data and became CODATA President for 1978-1982.



Mikhail Styrikovich

The final member of the first CODATA Bureau, Mikhail Styrikovich of the USSR, was more of an engineer than the other five. Through his knowledge of fluid dynamics and heat transfer he had made important contributions to the design of steam turbines. Styrikovich was keenly aware of the need for reliable data in the design of industrial processes of all types. He had a strong interest in international cooperation in science and technology and was probably instrumental in arranging for so many Soviet scientists to become active in CODATA, even at a time of deep political divisions.

The organizational structure approved by ICSU called for two classes of voting members, Union Members and National Members. All ICSU Unions were invited to join, without responsibility for paying dues. The National Members were assessed dues to provide the financial support of CODATA. There were also categories of Liaison Members and Co-opted Members. Each Member designated a Delegate to CODATA, who represented the Member in the General Assembly, the decision-making body. The Constitution, as approved by the ICSU General Assembly in 1966, gave the following assignments to CODATA:

- to ascertain, through the Unions and appropriate National bodies, what data-compiling work is going on and what the needs are;
- to achieve coordination among, and provide guidance for, data-compiling projects;

- to encourage support for data-compiling projects by appropriate private, governmental, and intergovernmental agencies;
- to encourage the use of internationally approved constants, units, and symbols and, when desirable, uniform editorial policy and procedures;
- to provide a compendium-directory-survey of continuing data-compiling projects and related work; and
- to encourage and coordinate research on new forms of preparing and distributing evaluated numerical data.

With this mission, CODATA came into being in 1966.

The 1966-1970 Period

Governance

The first task was to establish a Central Office to oversee the work of CODATA and implement the programs developed by the Bureau. The US National Academy of Sciences – National Research Council, which had previously set up an Office of Critical Tables (OCT) to monitor data compilation efforts in the United States, volunteered to provide temporary office space in Washington as well as the services of Guy Waddington, the Director of OCT. Waddington, a Canadian-born chemist who had studied the thermodynamic properties of petroleum products, thus became the first Executive Director of CODATA; he was, in fact, the person who coined the acronym CODATA for the full title of the organization, Committee on Data for Science and Technology of the International Council of Scientific

Unions. Three others were added to the initial staff of the Central Office: Christoph Schäfer, a German biologist from the Institut für Dokumentation in Frankfurt, Simone Kyropoulos from France, and Annigret Kundrat from the International Atomic Energy Agency in Vienna. In 1968 the Central Office was moved to Frankfurt, Germany, and Christoph Schäfer became Executive Director in July 1968. Martin Lewis from the UK joined the office as Scientific Editor.

The original six National Members of CODATA were the United States, United Kingdom, Germany (at that time the Federal Republic of Germany), France, Japan, and the Soviet Union. Canada joined as a National Member a short time later. Each National Member was sponsored by an appropriate organization in that country, which provided input to planning CODATA programs and was responsible for the dues to CODATA. The initial representatives of these members were:

Member	Representative	Primary scientific field	Sponsor
US	Frederick Rossini	Chemical thermodynamics	National Academy of Sciences
UK	Sir Gordon Sutherland	Molecular spectroscopy	Royal Society
Germany (BRD)	Wilhelm Klemm	Inorganic chemistry	Deutsche Forschungsgemeinschaft
France	Boris Vodar	High pressure physics	Centre national de la recherche scientifique
Japan	Masao Kotani	Atomic physics, biophysics	Science Council of Japan
USSR	Mikhail Styrikovich	Fluid dynamics, engineering	USSR Academy of Sciences
Canada	R. Norman Jones	Molecular spectroscopy	National Research Council of Canada

By 1968, ten ICSU Unions had become Union Members of CODATA:

Union

Representative

International Union of Biological Sciences (IUBS)	R. I. Currie
International Union of Pure and Applied Biophysics (IUPAB)	J. T. Edsal
International Union of Geodesy and Geophysics (IUGG)	G. Garland
International Geographical Union (IGU)	W. Garrison
International Union of Crystallography (IUCr)	Olga Kennard
International Union of Pure and Applied Chemistry (IUPAC)	Wilhelm Klemm
International Union of Theoretical and Applied Mechanics (IUTAM)	F. Odqvist
International Union of Geological Sciences (IUGS)	M. Roubault
International Astronomical Union (IAU)	Charlotte Sitterley
International Union of Pure and Applied Physics (IUPAP)	Boris Vodar

The Constitution also provided for Co-opted Members (later called Liaison Members), which included any type of international organization with an interest in scientific and technical data. The initial organizations in this category were:

ICSU Abstracting Board (IAB)
Federation of Astronomical and Geophysical Services (FAGS)
International Federation for Documentation (FID)
International Atomic Energy Agency (IAEA)
World Meteorological Organization (WMO)
Organization for Economic Cooperation and Development (OECD)

Finally, Guy Waddington was designated an individual Co-opted Member after completion of his service as Executive Director.

The CODATA Bureau met seven times between 1966 and 1970 in a variety of places - Paris, Moscow, Frankfurt, London, St. Andrews (Scotland), and Washington state (US). Among other decisions, it established a budget in US dollars and assigned a dues quota for each National Member. The largest contributor was the United States.

Each national sponsoring organization was required to raise its dues contribution in US dollars, not always easy in a period of frequent exchange controls and currency restrictions.

The Constitution called for an annual General Meeting, to be attended by the representatives (later called Delegates) of all classes of Members, where programmatic and budgetary decisions were to be made. The terminology was soon changed to Annual Meeting and later to General Assembly. The first two were simultaneous with Bureau meetings in 1966 and 1967, and the 3rd was held in 1968 in Arnoldshain, Germany, following the first CODATA International Conference. The 4th was held in Rome in 1969 (at which time Poland was admitted as the eighth National Member), and the 5th in 1970 in St. Andrews, Scotland, after the second CODATA Conference. At this 5th Annual Meeting the Constitution was revised to specify the membership of the Bureau as consisting of the four Officers, three members elected from CODATA representatives, and the immediate Past President and



First CODATA Conference, Arnoldshain, Germany, 1968

Executive Director as non-voting members. The number of National Members grew to 12 with the admission of Italy, German Democratic Republic, Israel, and The Netherlands.

CODATA Conferences

One of the early decisions of CODATA was to hold biennial conferences at which new developments in data compilation, critical evaluation, and dissemination mechanisms could be reported. These conferences were to stress both international and interdisciplinary information exchange. The *First International CODATA Conference on Generation, Collection, Evaluation, and Dissemination of Numerical Data for Science and Technology* was held in July 1968 at Arnoldshain, a resort town in the Taunus Mountains near Frankfurt am Main. This conference drew about 90 attendees from 12 countries (see Appendix F for attendance list). The program included summaries of data activities in the member Nations and member Unions; progress reports from individual data centers in thermodynamics, fluid properties, infrared spectroscopy, and other fields; and generic issues such as terminology, indexing, and classification. One session was devoted to the use of computers for data management, which was in an early stage in 1968 but was to expand rapidly. Panel discussions on future CODATA projects and priorities were also held.

The second CODATA Conference was hosted by the UK and held in September 1970 at St. Andrews, Scotland – a venue undoubtedly influenced by Sir Gordon Sutherland, the UK Delegate and Secretary-Treasurer of CODATA. In addition to progress reports of the physical science data activities that had been covered at the first Conference, there were talks on the feasibility of critical data evaluation in the geosciences and biological sciences. Attention was also given to data needs in industry and on ways to improve the presentation of data in the primary literature. An important feature of the conference was the coverage of computer use in data centers. Talks were given on computerized typesetting, visualization, and storage and retrieval of numerical data. Remote data retrieval was demonstrated by several groups, and terminals capable of handling complex notation and mathematical expressions were shown.

Scientific and Technical Activities

The 1968-1970 period saw the beginning of technical projects with specific objectives. These included preparation of a worldwide list of data centers and the establishment of Task Groups on Fundamental Constants, Key Values for Thermodynamics, Chemical Kinetics, and Computer Use. The first of these addressed the initial task specified in the Constitution, “To ascertain on a worldwide basis ... what work on critical compilation of evaluated data is being carried out in each country ...”. The work was undertaken by the Central Office in Frankfurt, which collected information from the National and Union Members, as well as other key individuals, in order to compile a list of continuing data centers and other projects. The coverage was restricted to physics, chemistry, materials science, and related engineering fields, and information was given on accessing the holdings of each center. The listing was published in September 1969 as a hard-cover book, *The International Compendium of Numerical Data Projects*. Updating and expanding this compilation continued as a CODATA activity for the next 25 years.

The founders of CODATA envisioned Task Groups as a mechanism for bringing together experts from different countries and different disciplines to attack specific problems of interest to the scientific and technical community. Funding for travel was to be provided by CODATA, and a prescribed output was expected. The first four Task Groups became active shortly after the 1968 CODATA Conference:

Fundamental Constants – Consistent values of the



Richard Cohen

highest possible accuracy for physical constants such as the speed of light in vacuum, mass of the electron, the Avogadro constant, etc. are important to all fields of science. This Task Group, initially chaired by E. Richard Cohen of the US, represented the first attempt to establish consistent values of these constants through an international consensus process. The work plan was to evaluate all relevant data from precision measurements and apply advanced statistical tools to arrive at the best possible

values of the individual constants. Their first report of recommended consistent values was published in 1973; subsequent activities of the Task Group will be described in the following sections.

Key Values for Thermodynamics – When accurate values of thermodynamic properties of individual substances are available, the well-established laws of thermodynamics allow one to predict what chemical processes can occur and how much energy is required or released in each process. This is an immensely powerful tool in both basic science and industrial applications. The value of a self-consistent set of thermodynamic parameters for important chemical substances was recognized by the 1930s. It was Frederick Rossini, in fact, who initiated the first efforts to compile such a data set while he worked at the National Bureau of Standards. It is not surprising that he felt this would be an appropriate project to demonstrate the value of CODATA. The resulting Task Group on Key Values, chaired by Stig Sunner of Sweden, held its first meeting at the Arnoldshain Conference in 1968, with initial members from Sweden, US, USSR, UK, and Germany. As will be described later, it worked in stages through the most important chemical elements, publishing interim reports as it went along.

Computer Use – The Computer Use group, chaired by Gordon Black of the UK, was established in 1968 with the task of “achieving maximum exchange of information about the methodology of handling data, including software, and to stimulate new work in this field”. The membership included both computer specialists and physical scientists who had recognized early the power of modern computers in handling their own data. Its first publication was *CODATA Bulletin No. 1, Automated Information Handling in Data Centers*, which appeared in 1969 and was followed by a second edition in 1971. These reports gave details about the hardware and software used in various data centers and thus became valuable resources for data centers that were beginning to explore the possibilities for automating their operations. The Task Group also organized the computer sessions and demonstrations at the Arnoldshain and St. Andrews conferences. It continued to make major contributions through the next decade.

Data for Chemical Kinetics – This group was established in 1969 with an initial goal of compiling information on data centers in the field of chemical kinetics and preparing a catalogue of evaluated compilations and critical reviews of kinetics data. Sidney Benson of the US was the first Chairman, and the first results were published as *CODATA Bulletin No. 3*. The Task Group expanded its mission and continued its work through the 1970s.

It is apparent from the description of the conferences and technical activities that CODATA focused almost exclusively on the physical sciences during its first five years, even though the charter given by ICSU was much broader. This can be attributed in part to the fact that the founders were chemists and physicists, but another factor was the prior existence within ICSU of the World Data Center program, which coordinated data collections in a variety of geoscience areas. CODATA eventually established good liaison, and several joint efforts, with this and other ICSU units, but it took some years to reach that point. As for the biosciences, the nature of the data was sufficiently different that there was only a limited history of systematic compilation of numerical data. This was to change dramatically with subsequent progress in molecular biology, and it will be described later how CODATA became a significant player in the field of bioscience databases.

Information Technology Activities

The founding fathers of CODATA came from an earlier generation and had no experience with and little knowledge of digital computers. In fact, some had reservations about the emphasis given to computers at the first CODATA Conferences. Nevertheless, they had the foresight to see that computers would become more significant as the volume and complexity of scientific data increased. The Task Group on Computer Use was the first step in addressing this issue. As time went on, CODATA became a pioneering leader in computer applications for managing scientific and technical data. In each succeeding decade, information technology played a larger role in CODATA activities.

Publications and Other Outputs

As CODATA started the various activities that have been described, attention had to be given to the means of disseminating the results. The *CODATA Newsletter* was started in 1968 as a vehicle for keeping the CODATA community – National Committees, Delegates, Task Group members, and interested data center personnel – informed. Each issue of the *Newsletter* also listed recent books, reports, and journals containing data compilations. These lists were potentially valuable to the broader scientific and technical users of data, and the National Committees were encouraged to distribute the *Newsletter* in their countries. However, it proved difficult to reach all the intended users with the limited resources available.



CODATA Bulletin Number 1

was later instituted. Reprinting articles from the *Bulletin* was encouraged; for example, issue No. 11, the first report on the CODATA Recommended Fundamental Constants, was reprinted in the *Journal of Physical and Chemical Reference Data*. The first CODATA publication, *The International Compendium of Numerical Data Projects*, was handled by a commercial publisher (Springer-Verlag), and other private publishers were utilized later for CODATA books of substantial length. In spite of the efforts of the Central Office, adequate dissemination of CODATA work to the scientific community was a continuing problem. The advent of the Internet has greatly facilitated this dissemination.

Summary

It is interesting to compare the path that CODATA embarked upon in its first five years of life with the original vision of the founders. This vision was articulated by Frederick Rossini in an article in the first *CODATA Newsletter*. He envisioned an “array of World Centers of Numerical Data for Science and Technology, covering each area of science, and appropriately tying in with the Central Office of CODATA as a hub.” The CODATA Central Office could become a repository for one set of all the data, or, at the very least, maintain a central directory of the holdings of all the centers, so that it could refer scientists to the center that would meet their needs. Another model, in his mind, was to have the Central Office process orders for data sets and direct the appropriate center to make the shipments. In this way CODATA might alleviate the currency conversion problems of scientists in developing countries (and occasionally in a developed country).

The reality was that CODATA lacked the financial resources and the authority to set up and operate such an ambitious network. What it did have was the ability to draw on the volunteer efforts of scientists who were committed to the goal of preparing the most reliable data sets possible. Thus it developed a program of action that facilitated the work of these scientists by providing a focal point for cooperative efforts and for information exchange, with the object of leveraging its limited financial and managerial resources to achieve maximum benefit.

To summarize its first five years, CODATA was successful in:

- reaching a membership of 12 countries, 10 ICSU Unions, and several other data-related organizations;
- establishing a governance structure and a Central Office that, after necessary adjustments, continues today;
- organizing two successful international and interdisciplinary conferences that were the first of a continuing series of biennial conferences;
- setting up the first scientific projects with a mandate to prepare accurate data sets to be recommended for international use;
- publishing a directory of worldwide data centers, and opening a forum on the use of computers to facilitate data compilation, evaluation, and dissemination.

In addition to these concrete achievements, the human interactions made possible by CODATA led to intangible but very significant benefits. The conferences and Task Group activities brought people together who might never have met otherwise. One example is worth noting. The work of the Task Group on Key Values for Thermodynamics was led by scientists from two institutions, the US National Bureau of Standards (NBS) and the USSR Institute for High Temperatures (IHT). This work started at the height of the Cold War, when tension was high and travel difficult. The rapport that developed among these individuals – Donald Wagman and William Evans of NBS and Lev Gurvich and Vadim Medvedev of IHT – continued for many years and led to many reciprocal visits. Interactions of this type, repeated many times within CODATA and other scientific organizations, made a significant contribution to the lowering of tensions during this difficult period.

The 1971-1980 Decade

Governance

CODATA entered the next decade with a sound governance structure in place and a nucleus of technical activities that were beginning to show results. It also had a new set of Officers, elected at the 1970 General Assembly, although the changes in leadership were relatively minor. Boris Vodar moved from Vice President to President and the two Vice President positions were occupied by Mikhail Styrikovich and Sir Gordon Sutherland. Wilhelm Klemm continued as Secretary-Treasurer. The new Bureau consisted of these four Officers plus three elected members: Lewis Branscomb of the US, R. Norman Jones of Canada, and Masao Kotani of Japan. Its first meeting was held in Paris in March 1971 and the 6th General Assembly took place in Washington, DC, in July 1971.

The 7th General Assembly was held in Le Creusot, France, in 1972, following the 3rd CODATA Conference. At the elections in Le Creusot, K. Egle of Germany replaced Wilhelm Klemm as Secretary-Treasurer; R. Norman Jones of Canada was elected Vice President; and Nicholas Kurti of the UK and Takehiko Shimanouchi of Japan joined the Bureau.



Norman Jones

analysis. When digital computers became available, he was one of the first spectroscopists to explore computer storage and retrieval of spectral data. From the beginning of CODATA, he pressed for recognition of the major role computers would assume in data management and dissemination. He was active in the Task Group on Computer Use and took a leading role in organizing the Freiburg Symposium on *Man-Machine Communication*

Norman Jones was a prominent Canadian spectroscopist, who pioneered the use of infrared spectroscopy for identifying chemical compounds. He was a leader in early efforts to collect and organize the images of infrared spectral traces for use in chemical

for *Scientific Data Handling*, which will be described below.

In 1973 CODATA faced a serious financial crisis that was precipitated by an upheaval in world currency markets. When the United States was forced to drop the fixed relation between the dollar and gold, the US dollar dropped sharply relative to European currencies. Since the CODATA dues assessments were set in dollars, while the Central Office expenses were incurred in German Marks, the budget became unbalanced. The 8th General Assembly, held in Stockholm in September 1973, decided to reduce the staffing level of the Central Office and to



**CODATA
Secretariat, Paris**

accept the offer of the French Government for office space in the former Hôtel de Noailles, where ICSU had its headquarters. Thus the Central Office was moved from Frankfurt to Paris at the beginning of 1974 and was renamed the CODATA Secretariat.

The 8th General Assembly gave the new Secretariat the mandate of implementing the decisions of CODATA and specified that it was to be headed by an Executive Secretary. The position was filled by Bertrand Dreyfus, a French physicist who had received part of his education in the United States. He had previously been in charge of information services at the European Space Research Organization and had a broad background in science and engineering. He engaged Phyllis Glaeser, an American



Phyllis Glaeser

with an engineering degree who had lived in Paris for many years, as an Administrative Assistant. The two of them provided outstanding leadership and support for CODATA activities until Dreyfus died in 1979, at which time Phyllis Glaeser replaced him as Executive Secretary.

In addition to reorganizing the Secretariat, the Stockholm General Assembly adopted a new Constitution that made

significant changes in the governance structure of CODATA. The Bureau was replaced by an Executive Committee consisting of the Officers plus five to eight ordinary members elected by the General Assembly from among the Delegates of the National and Union Members. The General Assembly would henceforth meet in alternate years, while the Executive Committee was to meet annually and would be responsible for transacting business between General Assemblies. The previous post of Secretary-Treasurer was split into separate Secretary General and Treasurer positions, with the Secretary General assuming responsibility for overseeing the activities of the Secretariat and generally coordinating the scientific program of CODATA. Terms of office for all elected positions were set at four years, with the possibility of one four-year renewal for all positions except President. Edgar Westrum (US) was elected as the first Secretary General and Nicholas Kurti (UK) became Treasurer.

This governance structure established by the 1973 Constitution remained largely in place through the remainder of the decade. The 9th General Assembly was held in Tsakhadzor, USSR, in 1974. The biennial schedule adopted in the new constitution was then followed, with the 10th General Assembly at Boulder (US) in 1976, the 11th at Santa Flavia (Italy) in 1978, and the 12th in 1980 at Kyoto (Japan). Each of these authorized new CODATA scientific activities which will be described below.



Paul Melchior

The elections at the 9th General Assembly in 1974 brought a new generation into CODATA leadership. The new President was Paul Melchior of Belgium who had been the Delegate to CODATA of the International Union of Geodesy and Geophysics. He was a distinguished astronomer, a personal friend of the King of Belgium, and later served as the Director of the Royal Observatory of Belgium. His special scientific interest was earth tides, and he became the acknowledged world expert on this subject. As President he helped bring a new perspective to CODATA, whose programs, as has already been noted, were previously dominated by chemists and physicists.

Thomas Plebanski, a chemist from Poland who had established a program for standard reference materials in his country, was elected Vice President and served for the next four years. Edgar Westrum was re-elected and continued to serve as Secretary General for the rest of the decade. His field was chemical thermodynamics, and he was a strong supporter of the efforts to compile and



Edgar Westrum

evaluate thermodynamic data. He was instrumental in setting up the publication programs of CODATA, and he established close relations with UNESCO and its UNISIST program, leading to a series of jointly supported training courses in data management for developing countries. Westrum provided a strong guiding hand at the Secretariat and was heavily involved with the *Newsletter*, as well as serving as editor of the *Bulletin*.



Nicholas Kurti

As Treasurer from 1973 to 1980, Nicholas Kurti influenced many CODATA activities. He was a Hungarian-born physicist who joined the exodus of Hungarian scientists in the 1930s and became a leading researcher in low-temperature physics at Oxford. In the 1950s his laboratory achieved the lowest temperatures yet recorded. He served a term as Vice President of the Royal Society, and an important prize, the Nicholas Kurti European Prize for Physics, has been established in his memory. Kurti was a very colourful character, an inspiring teacher, and an expert on the science of cooking – a topic on which he often gave TV demonstrations. His extensive international contacts were important as CODATA expanded.



Takehiko Shimanouchi

At the 10th General Assembly in 1976 a second Vice President was added, Takehiko Shimanouchi of Japan. He was a molecular spectroscopist who made important contributions to the use of infrared and Raman spectroscopy in elucidating the structure of small molecules and polymers. He also made pioneering contributions to codifying

and evaluating spectroscopic data by publishing a series of definitive papers on fundamental vibrational frequencies in small molecules. Shimanouchi resigned in 1978 when his countryman, Masao Kotani, was nominated for President. He continued to be active in CODATA up to his untimely death while he was organizing the program for the 7th CODATA Conference.

Masao Kotani was one of the six founders of CODATA, and he was elected President at the 11th General Assembly in 1978. Kotani had an unusually broad scientific background. He helped introduce quantum mechanical techniques into Japan in the 1930s, applying them to atomic collisions and molecular structure calculations. He also worked on microwave circuits and components, later turning his attention to biophysics, where he applied quantum chemistry calculations to biological molecules. As President of the Tokyo Science University, he had a strong influence on the education of the next generation of Japanese scientists. He later received the Order of Cultural Merit from the Emperor of Japan. His tenure as CODATA President coincided with the first stages of CODATA attention to bioscience data.

Also elected at the 11th General Assembly was V. V. Sytchev of the USSR as Vice President. He was a specialist in the properties of fluids and a Professor at the Moscow Power Engineering Institute. Sytchev was the organizer of the 4th CODATA Conference in Armenia. Edgar Westrum continued as Secretary General and Nicholas Kurti as Treasurer.



Jacques-Émile Dubois

had realized very early that digital computers would revolutionize the method of identifying and classifying chemical compounds. He created the DARC system for representing the geometric structure of a molecule by a digital record that allows searching for a desired structure in a database of millions of compounds. DARC has been

heavily used in operational search systems for both research reports and patents. Dubois served at various times as a high-level advisor to several agencies of the French Government and to industrial concerns. He was to play a major role in CODATA for the next 20 years, chairing task groups addressing computer innovations and serving as President from 1994-1998.



David Watson

instrumentation and increased research on biochemical molecules flooded the literature with new crystallographic data. Watson chaired the Commission on Crystallographic Data of the International Union of Crystallography and in 1972 became the IUCr Delegate to CODATA. He was a leader in CODATA educational projects before being elected Treasurer, and he exercised a steady hand in managing CODATA finances in the 1980s.



Officers elected in 1980 – David Watson, Masao Kotani, Edgar Westrum, V. V. Sytchev, Jacques Dubois

CODATA Conferences

The 3rd International Conference was held at Le Creusot, in the Burgundy region of France, in June 1972. One objective of the Conference was to explore the prospects for new projects in the broadened scope of CODATA. Discussions with ICSU had clarified that the scope of CODATA was to include not only quantitative information on the properties and behavior of matter, but also other experimental and observational data. The Conference program thus included sessions on data needs in:

- Earth and Atmospheric Sciences,
- Biological Sciences,
- Astronomy and Astrophysics, and
- Engineering Sciences.

Other sessions dealt with cross-cutting issues such as computerized data handling and data publication issues. Demonstrations of remote access to data banks were a popular feature of the Conference. In contrast to the first two Conferences, the proceedings of the Third Conference were published by CODATA itself.

The USSR was host for the 4th CODATA Conference, which was held in the Tsakhadzor Olympic Village in Armenia in 1974. The program included several sessions on thermodynamic and thermophysical data compilation, an area of great activity in the Soviet Union. As in the previous Conference, emphasis was given to data in the biosciences, geosciences, and astronomical sciences. One interesting feature was a talk by Ronald Wigington, using motion pictures, on computer-generated graphics as a means of displaying data on Brownian motion and protein folding. The Conference was attended by about 130 scientists from 13 countries.

The United States hosted the 5th CODATA Conference in Boulder, Colorado, a few days before the bicentennial celebration of American independence in July 1976. The attendance of almost 400 was considerably larger than that at previous conferences, and the program was more extensive. Much of the program focused on pressing world problems such as energy conservation and environmental protection. Talks included water quality data, meteorological data, and reactor safety codes. One well-attended session described data collections on the

volatile constituents of wines. As in the two previous Conferences, new developments in computerized databases received heavy emphasis.

The 6th Conference was organized by the Italian National Committee and held at Santa Flavia in Sicily in 1978. The program showed continued emphasis on data relevant to environmental pollution and energy issues. In keeping with the location of the Conference, there were presentations on data needs for prediction of earthquakes and volcanic eruptions. Other talks foreshadowed future changes in the broader areas of information technology, such as the computerization of research journals.

The 7th CODATA Conference, the final of the decade, was held in Kyoto, Japan, in October 1980, with the theme *Role of Data in a Dynamic World*. The increasing importance of reliable data in attacking major world problems was emphasized in the keynote address given by Phillip Abelson, Editor of *Science*. More sophisticated computer techniques for chemical structure searching, pattern recognition, and data compression were described. Bioscience databases received greater emphasis, reflecting the growth in digitized collections of data in microbiology, protein structure, and drug properties. It is also interesting that talks on copyright and intellectual property issues, as well as the economics of online data dissemination, were given for the first time at a CODATA Conference.

Scientific and Technical Activities

Five technical projects had been initiated by the end of the previous decade: three Task Groups dealing with specific disciplinary areas, one that addressed computer applications, and one project aimed at identifying existing data centers throughout the world. A number of new projects were established in the 1971-1980 period. The five original projects will be covered first.

Fundamental Constants – Chaired initially by E. Richard Cohen (later by Terry J. Quinn, Barry N. Taylor, Peter J. Mohr, and Barry M. Wood), the Task Group has included, over its 45 year history, members from all the world laboratories engaged in high-precision physical measurements. Starting in 1969 the group made a study of all relevant experimental data and theoretical constraints and arrived at a recommended set of

fundamental constants in 1973. The set was published in the newly established *Journal of Physical and Chemical Reference Data* as well as in *CODATA Bulletin No. 11*. The participation of a broad range of knowledgeable experts in this effort helped to ensure the near universal acceptance of the Task Group's recommended set of values by the international scientific community.

Important experimental and theoretical work in the precision measurement-fundamental constants field continued apace and the Task Group continued to meet and analyze the new data, leading to the publication of a new and significantly more accurate set of recommended values in 1986. The release of the 1986 CODATA recommendations was reported in the *New York Times*, which commented "The revised list of constants ... sets the seal of international acceptance on improved measurements made during the last 13 years...[T]he publication marks some important milestones".

Key Values for Thermodynamics – As mentioned earlier, this group was first chaired by Stig Sunner of Sweden, and most of the data evaluation was done at the National Bureau of Standards in Washington and the Institute for High Temperatures in Moscow. The first definitive set of recommended values was published in 1973 as *CODATA Bulletin No. 10*. Several interim updates were published later, leading to a new comprehensive set in 1978 (*CODATA Bulletin No. 28*). The group continued to work for another decade under the chairmanship of John Cox of the UK, eventually extending the coverage to 151 key chemical species. The final results, including extensive documentation of the data analysis, were published as *CODATA Key Values for Thermodynamics*, J. D. Cox, D. D. Wagman, and V. A. Medvedev, eds., Hemisphere Publishing Corp., 1989. These values have been widely adopted by the thermodynamics research community as the basis for calculating thermodynamic tables and predicting properties of chemical reactions.

Chemical Kinetics – The first output of the group was the 1974 publication of *CODATA Bulletin No. 13, The Presentation of Chemical Kinetics Data in the Primary Literature*, which addressed the need for authors of experimental papers to give full information on their measurement methods and to express their data in a

systematic format. As the concern about depletion of the earth's ozone layer grew, the Task Group was reconstituted in 1977 with a mandate to produce a set of critically evaluated rate parameters for reactions pertaining to stratospheric chemistry. The new Chairman was J. Alistair Kerr of the UK, and members came from the US, Germany, and the USSR. A first set of recommended data was published in 1979 as *CODATA Bulletin No. 33*; full documentation appeared shortly afterwards in the *Journal of Physical and Chemical Reference Data*. The Task Group then expanded its attention to other chemical reactions relevant to atmospheric pollution and, later, to global warming. After publication of a supplement to the first data set in 1982, CODATA reached agreement with the International Union of Pure and Applied Chemistry to continue the work under IUPAC auspices.

CODATA's role in atmospheric chemistry data was extremely influential. It launched the first effort to provide an internationally accepted database for use by atmospheric modelers and thus demonstrated that it could move quickly to set up a much-needed project that eventually morphed into large-scale continuing programs for data collection and analysis supported by major governments and international organizations such as the Intergovernmental Panel on Climate Change (IPCC).

Computer Use – In 1973 the Task Group on Computer Use organized a week-long symposium in Freiburg, Germany, on the subject *Man-Machine Communication for Scientific Data Handling*. The symposium brought together more than 100 scientists from 14 countries representing fields ranging from seismology to weather forecasting. The state of the art in hardware and software was reviewed, and a wide range of applications to the collection, management, and dissemination of scientific and technical data were described. Special emphasis was given to data sharing through telecommunication networks, a practice still in its infancy. The conclusion of the symposium was that computer technology had reached the stage where no significant technical barriers existed to the widespread use of automated techniques for all aspects of data management, and that the remaining issues were primarily economic and legal in nature.

The Freiburg symposium was one of CODATA's most significant accomplishments in the decade. It brought together scientists in different fields who were using new information technology to solve specialized problems of data management but who would not normally have a means of communicating with each other. The resulting cross-fertilization undoubtedly contributed to the rapid expansion of computer applications in the 1970s.

CODATA Directory of Data Sources for Science and Technology – After completion of *The International Compendium of Numerical Data Projects* in 1969, it became clear that a single publication covering projects in all disciplines and countries would be very hard to maintain. Accordingly, a decision was made to start a series of directories of data sources in specific scientific disciplines. Scientists working in each field were to be enlisted to gather current information on active data centers and other data sources. During the next decade, 12 chapters in this series were published as CODATA Bulletins. Although distribution was not as wide as hoped, these directories did prove valuable to librarians and technical information specialists. With the advent of Internet search engines, much more efficient tools for locating data sources became available, but the CODATA Directories provided a useful service in their time. The fields covered were:

Chapter	Discipline
1	Crystallography (1977)
2	Hydrology (1979)
3	Astronomy (1980)
4	Zoology (1980)
5	Seismology (1981)
6	Chemical Kinetics (1981)
7	Nuclear and Elementary Particle Physics (1982)
8	Atomic and Molecular Spectroscopy (1982)
9	Snow and Ice (1984)
10	Geodesy (1984)
11	Chemical Thermodynamics (1984)
12	Nutrition (1984)

During the 1971-1980 decade a number of new Task Groups and other projects were established. Some of these completed their work in this decade, while others continued into the 1980s. The most important are listed below:

Accessibility and Dissemination of Data – This important Task Group, often referred to as ADD, was established in 1972 under the impetus of Masao Kotani and with joint support by UNESCO. Its lengthy report, *Study on the Problems of Accessibility and Dissemination of Data for Science and Technology*, was published in 1975 as *CODATA Bulletin No. 16*. The report proposed a detailed classification scheme for the various categories of data and described the problems of collection, evaluation, and dissemination of data in each category. It went on to outline a rather idealistic model for an integrated world system of scientific data management and concluded with recommendations for action by governments, UNESCO, CODATA, and other participants. Although few of these ambitious proposals could be implemented, the ADD Task Group stimulated several new CODATA projects.



ADD Task Group, 1973 – Masao Kotani (presiding), unknown, unknown, I. Eliezer, Steve Rossmassler

Training Courses in the Handling of Experimental Data – This Task Group, an outgrowth of the ADD report, was set up to organize training courses to inform librarians and information specialists in developing countries about sources of data relevant to the needs of their clientele. With financial support from UNESCO, CODATA conducted training courses in Yugoslavia, Poland, Japan (with attendance of librarians from other Asiatic countries), Brazil, and elsewhere. David Watson



**CODATA-UNESCO Training Course,
Tsukuba, Japan, 1980**

of the UK, Antoinette David of France, and Edgar Westrum of the US were especially active in these courses. The reception from the attendees was always positive, and it seems clear that CODATA made a measurable contribution to helping scientists in developing countries find needed data in the era before the Internet. CODATA and UNESCO continued to conduct these training courses through the 1980s; in 1986 David Watson wrote a very detailed set of guidelines for the organization of such training courses and workshops.

Presentation of Data in the Primary Literature –

Another issue discussed in the ADD report was the need for scientists who produce data to report full details of their measurements, in order that evaluators, as well as the ultimate users, can understand the meaning of the data without ambiguity. CODATA established a Task Group, with Edgar Westrum as Chairman, to address this subject in 1972. Its first publication was *CODATA Bulletin No. 9, Guide to Presentation in the Primary Literature of Numerical Data Derived from Experiments*, which covered general principles. This was followed by several publications focused on specific types of data, including:

- *Biochemical Equilibrium Data* (No. 20, 1976)
- *Biologist's Guide* (No. 25, 1977)
- *Observational Data in the Geosciences* (No. 32, 1979)
- *Calorimetric Measurements on Cellular Systems* (No. 44, 1981)
- *Astronomical Data* (No. 46, 1982)

Transport Properties – This Task Group was organized by Yeram Touloukian in 1972 with the aim of producing key data sets analogous to those of the Key Values for Thermodynamics group. It was reconstituted in 1974 as the Task Group on **Thermophysical Properties of Solids**, with Merrill Mingos (US) as Chairman. Its main outputs were two publications, *Thermophysical Properties of Some Key Solids* (No. 59, 1985) and *Evaluation of Thermophysical Property Measurement Methods and Standard Reference Materials* (No. 60, 1986).

Data for the Chemical Industry – This Task Group began work in 1976 under the leadership of Arnold Bondi (US). In addition to conducting a survey of chemical engineers on types of property data and classes of chemical compounds that were of highest priority, it focused on data estimation procedures needed in the industry. The group published a guide to physical property correlation techniques and estimation procedures in 1978 (*CODATA Bulletin No. 30*).

Internationalization and Systematization of Thermodynamic Tables

Several CODATA leaders felt that, in addition to providing the key values to which comprehensive thermodynamic tables are tied, CODATA might have a useful role in promoting generation of these more extensive tables. A Task Group on this topic was established in 1976, chaired by Howard White of the US and including members from the UK, USSR, US, France, and Germany. Its charge was to reach international agreement on conventions, notation, standard states, and other features that would facilitate the preparation of compatible and consistent tables. It published *A Systematic Approach to the Preparation of Thermodynamic Tables* in 1982 as *CODATA Bulletin No. 47*. A successor Task Group on Chemical Thermodynamic Tables was established in the same year to develop prototype tables based on these recommendations.

Data Flagging and Tagging – A problem faced by data centers, as well as ultimate users of data, was the identification of research papers that report new measurements. Abstracts often fail to mention all the data contained in a paper, and useful data may be overlooked. CODATA and the ICSU Abstracting Board set up a Joint Working Group to address this issue in 1974. In the resulting publication, *CODATA Bulletin No. 19*, the group

surveyed existing practices in various disciplines and made recommendations for improving the compatibility of the different approaches. However, the utility of data flagging soon diminished as full-text searching of digital versions of journal articles became possible.

World Data Referral Center – One of the recommendations of the ADD report was that CODATA undertake an operational role as a world focus for referring users to reliable data sources. Such a role had also been envisioned by the CODATA founders, as described previously. In 1977 a World Data Referral Center was established with financial backing from UNESCO and other sources. The concept was to have a dedicated staff in the ICSU building who would answer questions from users throughout the world, providing the requested data or, if that was not possible, referring the user to a source that could provide it. Unfortunately, the task proved far too ambitious for the resources available, and the project ended after a few years.

Publications and Other Outputs

During this decade the *CODATA Bulletin* became the primary outlet for publishing the results of CODATA projects. Thirty-seven issues were published between 1971 and 1980. Sixteen issues of the *Newsletter* also appeared. Proceedings volumes were published for the CODATA Conferences at Le Creusot, Tsakhadzor, Boulder, Santa Flavia, and Kyoto. The first was published and distributed by CODATA, the others by Pergamon Press. Several volumes of selected papers from the Conferences also appeared, covering topics such as energy-related data and man-made hazards. The following book should also be mentioned: *Data Handling for Science and Technology - An Overview and Sourcebook*, S. A. Rossmassler and D. G. Watson, eds., Elsevier (North-Holland), 1980. This sourcebook was one of the first, if not the first, to discuss computer handling of scientific and technical data in a detailed way.

Summary

CODATA could claim a number of successes, plus a few failures, during this decade. The most significant achievements included:

- The CODATA International Conferences were established as biennial events that drew participants from a range of disciplines and countries.
- The Freiburg Symposium on *Man-Machine Communication for Scientific Data Handling* signaled a leadership role for CODATA in promoting the application of modern information technology to scientific data management.
- CODATA became the recognized international authority for providing definitive values of the fundamental constants, key thermodynamic data, and chemical kinetic data needed for atmospheric modeling.
- CODATA began to broaden its disciplinary scope by bringing data experts in the geosciences and biosciences into its biennial Conferences and its directory projects, setting the stage for more extensive collaboration that was to occur in the 1980s.
- Joint projects between CODATA and UNESCO were established with a focus on data needs in developing countries.
- CODATA expanded its membership to 16 Nations plus 17 Unions and other ICSU bodies.

The 1981-1990 Decade

Governance

The 13th General Assembly was held at Jachranka, Poland in 1982, during a difficult political period. As a protest against suppression of free expression in Poland, the United States and other countries placed travel restrictions that reduced the attendance. Nevertheless the CODATA Conference and General Assembly did take place. In addition to approving five new Task Groups, the Assembly elected new Officers and Executive Committee members.



**William
Hutchison**

The new President was William Hutchison, who had previously served as International Union of Geological Sciences (IUGS) Delegate to CODATA and as a member of the Executive Committee. Born in Scotland, he had become a leading geologist in Canada and at this time served as Assistant Deputy Minister for Earth Sciences in the Canadian Government. He had extensive international experience, including many years of interactions with UNESCO and a term as Secretary General of IUGS. Hutchison was a strong leader and an eloquent spokesman for international science.



David Lide

His research career was in molecular spectroscopy, and since 1969 he had directed the Standard Reference Data program at the National Bureau of Standards. He was the Founding Editor of the *Journal of Physical and Chemical Reference Data*, a journal dedicated to publishing critically evaluated data in the physical sciences. He was also active in IUPAC, where he was serving at this time as President of the Physical Chemistry Division.

David Lide of the United States was elected Secretary General. He had been involved with CODATA since the first Conference in 1968, serving as US Delegate, Executive Committee member, and Associate Editor of both the *Newsletter* and *Bulletin*.

The first decision of President Hutchison was to hold a strategic planning meeting to review past activities and outline a course for CODATA to take in the future. The meeting was held at La Gaillarde, France, in June 1983 and was attended by the Executive Committee and invited experts. The group held extensive discussions on changes in the scientific world in the almost 20 years since CODATA was conceived by its founders and on how CODATA should change its directions to meet these challenges. Consensus was reached on several points:

- Emphasis should be shifted in the direction of work on generic data structures and data management issues of interest to all disciplines, rather than a focus on specific data sets for a single discipline.
- In view of the rapid advances in molecular biology and the resulting deluge of data, higher priority should be given to bioscience data.
- Advances in computer technology provide a major opportunity for CODATA leadership in applying modern information technology to data management and dissemination.
- CODATA should promote stronger ties to the Unions and other ICSU bodies.
- Every effort should be made to propagate CODATA's outputs to the scientific and technical community.

The conclusions of the planning conference provided a framework for choosing CODATA activities during the following decade.

There was no need for the 14th General Assembly to elect new Officers, but a number of important actions were taken when it met in Jerusalem in 1984. The conclusions of the La Gaillarde planning meeting were endorsed and the Constitution was amended to reflect the restructuring recommended by the planning group. Under the new structure, Task Groups were now required to have a well-defined goal and time scale, and term limits were set for their Chairs. A new category of Working Groups was created to undertake short-term studies that might lead

to the formation of a Task Group or some other activity. The term Commission was to be applied to a group set up to monitor a broad range of activities and advise the Executive Committee. Finally, a new dues structure was approved and terms of payment were clarified.

Another important action at the 14th General Assembly was the admission of the People's Republic of China as a National Member. China was to become a leading participant in CODATA in the future.



Andrzej Bylicki

At the 15th General Assembly in 1986, held in Ottawa, David Lide was elected President, Jacques Dubois continued his term as Vice President, and David Watson continued as Treasurer. There were two new Officers; Andrzej Bylicki was elected second Vice President and Alain Bussard became Secretary General. Bylicki was a Polish physical chemist with a special interest in phase equilibria and properties of organic compounds of industrial importance. He spent most of his career at the Institute of Physical Chemistry of the Polish Academy of Sciences. He was a champion of Polish participation in CODATA and continued his own involvement for many years.



Alain Bussard

Alain Bussard was a distinguished French biologist whose career at the Pasteur Institute began in 1945. He worked in cellular biochemistry and antibody synthesis and became head of the Department of Immunology. He was one of the first to obtain continuously growing hybridoma cell lines. As Secretary General he oversaw the expansion of CODATA activities in the biosciences and was especially active in the Hybridoma Data Bank project.



James Crease

Only one new officer, Treasurer, was elected at the 16th General Assembly, held in Karlsruhe, Germany, in 1988. This was James Crease, an oceanographer from the UK. He had worked in British Government laboratories and later became Adjunct Professor of



Officers Meeting, Moscow, 1987

Oceanography at the University of Delaware, where he designed databases for tracking oceanographic research vessels and managing the data they collected.

The 17th General Assembly was held in Columbus, Ohio, in 1990. One of its most significant actions was reaching agreement on the admission of the Chinese Academy located in Taipei as a National Member in parallel with the Chinese Academy located in Beijing, which had become a member in 1984. The inclusion of both Academies in the CODATA family brought in a number of scientists who were to make important contributions to future CODATA projects.



David Abir

Except for Treasurer, where James Crease continued, a complete slate of new Officers was elected in 1990. David Abir of Israel, a Professor at the Technion, was elected President. He was an aeronautical engineer with special interest in properties of engineering materials and fluid dynamics. Lev Gurvich of the USSR became Vice President. Gurvich was a thermodynamics specialist who had been active in



Lev Gurvich

CODATA since the first Conference in 1968. He was a prime mover in the Key Values for Thermodynamics Task Group, and he led the conversion of the large database at the Institute for High Temperatures in Moscow to digital form. Gordon Wood, who was elected Secretary General, was a Canadian low-temperature physicist who developed Canada's 'Josephson Volt' – the Canadian Standard Volt linked to



Gordon Wood

He had been active in CODATA since 1980, including resurrecting the Canadian National Committee for CODATA and organizing the 1986 CODATA Conference in Ottawa.

CODATA Conferences

The 8th International CODATA Conference was originally scheduled for Kozubnik, Poland, in October 1981, but political instability in Poland at that time forced a change of venue to Jachranka, a suburb of Warsaw. The theme of the Conference was *Data on Natural Resources – Their Use for the Development of Society*. The program showed a focus on raw materials and fossil fuels, emphasizing the importance of reliable data in their efficient utilization. In spite of logistic difficulties and travel restrictions, a successful Conference was conducted. More than 160 scientists from 22 countries participated.

Jerusalem was the site of the 9th Conference, held in 1984. It attracted over 250 participants from 20 countries. The distribution of papers over scientific disciplines was broader than at previous Conferences. There were active sessions on nucleic acid and protein sequence analysis, oceanographic and ecological data collection, data from nuclear and particle physics, and the usual topics in chemistry. As in the two previous Conferences, computer applications were highlighted. In spite of the touristic attractions of the site, the sessions were well attended. Unfortunately, few scientists from Eastern bloc countries were able to attend. Thus the geographical distribution of attendees in the 9th Conference was a mirror image of the 8th.

The 10th Conference in Ottawa, Canada, in 1986 achieved a more balanced geographical distribution. The theme of the 10th Conference, *Computer Handling and Dissemination of Data*, reflected the dramatic growth of the role of the computer over the 20 years of CODATA's existence. The program was designed to bring out the common features of data management and database design in the various scientific disciplines. Particular

emphasis was given to the new role of microcomputers, and several talks described plans for converting print sources of data, such as the *Beilstein Handbook of Organic Chemistry*, to digital databases. Over 300 people attended.

In 1988 CODATA returned to Germany for the first time since its first Conference, holding the 11th CODATA Conference in Karlsruhe with the theme *Scientific and Technical Data in a New Era*. Close to 200 people from 27 countries attended. The subject coverage was broad, with particular emphasis given to data relevant to biotechnology and to industrial needs. The topic of global climate change appeared on a CODATA Conference program for the first time. Another first was a session on legal and social aspects, which included papers addressing copyright, liability, and privacy issues. These issues were to receive increased attention from CODATA in the future.

The 11th CODATA Conference, held in Columbus, Ohio, in 1990, was the second held in the United States. Its theme was *Data for Discovery*, and it attracted 365 participants from 29 countries (more than the 26 National Members of CODATA at the time). The opening talk by Kenneth Wilson, Nobel Laureate in Physics in 1982, described the coming revolution in large scale computing and its effect on the future of science. The global change program and its data needs now received major billing. Bioscience data were again emphasized, with talks on the human genome project, sequence databases, and pharmacological data. Early attempts to extract new knowledge from combinations of numerical databases (now called “mashups”) were described. A well-attended final session reviewed the history of data compilations, CODATA's impact over the previous 25 years, and future CODATA directions.

Proceedings volumes covering at least the major papers were published for all the Conferences in this decade.

Scientific and Technical Activities

The most significant aspect of CODATA's program in this decade was the expansion into the biosciences and geosciences. Another important trend was the increasing attention to modern information technology and its application to data management and dissemination. The following selection of new Task Groups authorized in the 1981-1990 period illustrates both these features.

Hybridoma Data Bank – The development of the hybridoma technique for the production of monoclonal antibodies in 1975 stimulated a rapid expansion of antibody research. This Task Group was established in 1982 in collaboration with the International Union of Immunological Societies. Its goals were to foster development of an international information system for identification and standardization of the nomenclature describing these reagents and to set



**Lois Blaine,
Director of
Hybridoma
Databank**

up a data bank to manage and disseminate hybridoma data. It was chaired by Alain Bussard of the Pasteur Institute and initially included members from the US, UK, Norway, and Japan. Lois Blaine of the US, who had extensive experience in developing medical and environmental databases, was engaged as Manager of the Hybridoma Data Bank. Micah Krichevsky and Bernard Janicki, both from the US National Institutes of Health, were particularly active in designing the data bank and obtaining financial support; Janicki became Chairman in 1986.

Under the CODATA umbrella, three identical copies of the Hybridoma Data Bank were housed on computers at the US National Institutes of Health (NIH), Japan's Institute for Physical and Chemical Research (RIKEN), and the Faculté de Médecine de Nice, France (later moved to Dialog's DataStar system by the European Node). Central operations were located at the American Type Culture Collection (ATCC) and directed by the Task Group. Chairman Alain Bussard was assisted by respected immunologists from the US, Switzerland, Germany, the World Health Organization, UK, Norway, and Japan.

It was estimated that over 10,000 laboratories were developing hybridomas when the data bank was initiated. Data were gathered through direct solicitation from the research community and by data extraction from catalogs, abstracts of scientific proceedings, and journals. Bulk entry of large data collections was also accomplished. All data entries were translated into controlled terminology to ensure interoperability and data sharing by laboratories throughout the world.

CODATA exhibited great foresight in recognizing the value of centralized access to data on biological products that have become a multibillion dollar industry and have inestimable value in biomedical research and therapeutics. Although the Hybridoma Data Bank was an extremely useful tool for biologists for about ten years, the exponential growth of hybridoma and monoclonal antibody development rendered it impossible financially and logistically to provide a comprehensive centralized source of information on these widely used reagents. At present there are numerous sources for identification of specific immunoreagents that can be accessed with Internet search tools.

Protein Sequence Data Banks – CODATA's initial involvement in this area was led by Borivoj (Boja) Keil of the Pasteur Institute. The Task Group was established in 1984 with Keil as Chairman and Akira Tsugita of Japan as Secretary; its first publication was a *Directory of Protein and Nucleic Acid Sequence Data Sources* (CODATA Bulletin No. 65, 1987). In the same year, it joined with other groups in organizing the *First CODATA Workshop on Nucleic Acid and Protein Sequencing Data at Gaithersburg, Maryland*.



Boja Keil

Three influential CODATA books were stimulated by the Task Group:

- *Biomolecular Data: A Resource in Transition*, Rita R. Colwell, ed., Oxford, 1989;
- *Computational Molecular Biology: Sources and Methods for Sequence Analysis*, Arthur M. Lesk, ed., Oxford, 1988; and
- *Information Integration for Biological Macromolecules*, Winona Barker, ed., CODATA Bulletin Vol. 23, No. 4, 1991.



**Biological
Macromolecules
Task Group –
Josh Lederberg, Nobel
Laureate, seated center**

In the early 1990s the group was reconstituted under the name **Biological Macromolecules**, with Akira Tsugita of Japan as Chairman, and it continued to have an active role in disseminating information on database activities, nomenclature, and standardization

efforts. It has also served a valuable function as a communication channel among the many groups archiving sequence data.

Microbial Strain Data Network – Biotechnology processes commonly involve the use of a single strain of microorganism with carefully defined properties. In response to a suggestion made at a meeting of the International Congress of Microbiology, CODATA joined two other components of ICSU, the World Federation of Culture Collections (WFCC) and the International Union of Microbiological Societies (IUMS), in creating the Microbial Strain Data Network (MSDN). CODATA established a Task Group on the MSDN to set policy and oversee the operation of a database describing the individual properties of microbial strains. The Task Group realized the impracticality of assembling detailed strain data from many laboratory strain collections in one repository. Local databases of strain data are best curated and maintained by the producers of the data. Instead, the model chosen was to establish a Central Directory of the data elements utilized by each of the collections to describe their strains. The Directory established codes for data elements describing all cellular microorganisms such as bacteria, fungi, and protozoa. The individual element descriptors were standardized by use of open-ended common nomenclature (e.g., dextrose coded as D-glucose). Thus, a unique number denoted each concept containing the term “D-glucose”. The use of a numerical index allows translation of the attached concept into other languages; local assignment of synonyms, abbreviations, and alphanumeric searching; and sorting across databases from disparate data sources.

MSDN conducted workshops on use of the central directory and efficient coding of strain data in venues around the world, especially in developing nations. Establishing or improving searchable databases was the emphasis of each of the workshops. With longer time available (usually five days), a full training course on coding strain data and computer usage of the resulting database was possible.

Standardized Terminology for Access to Biological Data Banks – This Commission was established in 1988 to promote improved standardization of terminology in the growing number of bioscience data banks. Its designation as a CODATA Commission, rather than a Task Group, signified that its mission was to catalyze ongoing efforts in the area of biological data interoperability. Therefore, it had broad membership representing different areas of biology and it developed close interactions with many ICSU Unions and other organizations concerned with bioscience data. It was chaired first by Lois Blaine and later by Frank Bisby.

Recognizing that the scope of biological data is enormous but that key points of entry are often the name and characteristics of an organism, the Commission focused attention on biological collections and their growing body of data. Obtaining cooperation in establishing metadata standards to facilitate data exchange among these collections was the Commission’s major goal. A workshop entitled “Building blocks for common access to biological collections”, held at the National Center for Ecological Analysis and Synthesis (NCEAS) in Santa Barbara, California, was attended by computer scientists and biologists from Europe, Latin America, Japan, and the US. The purpose of this meeting was two-fold:

- to develop a common data specification for biological collection units (including simple field observations that did not produce preserved and cataloged specimens); and
- to develop common software architecture to support distributed queries across collection databases.

All of the world’s biological collections contain a number of data items including specimen specific (e.g. taxon, date, altitude, sex) and collection specific (e.g. holding

institution) elements. The set of elements used varies from collection to collection, and there were no widely adopted standards for common sets of elements. Clearly, a reconciled set of element names for scientists and curators to use was needed. A design goal of the data specification was to be both comprehensive and general, to include a broad array of concepts that might be available in a collection database, but to mandate only the bare minimum of elements required to make the specification functional. The world's collection databases utilized myriad formats and access technologies. Many of these were primitive, hard to use, platform specific, and scaled poorly. Furthermore, almost all of the existing systems were incompatible with each other. Rather than encourage database providers to take on the burden of redesigning/rebuilding existing databases, a software architecture that defined a system of "gateways" to allow existing, otherwise incompatible databases, to be interrogated, as well as "portals" to access the gateways, was required for interoperability.

Working closely with the Species2000, the Global Biodiversity Information Facility (GBIF), and the Taxonomic Data Working Group (TDWG), the work of the CODATA Commission has evolved from the time of the defining workshop in 2001. A Working Group entitled Biological Collection Data Access (renamed to Observation & Specimen Record) was approved by CODATA in 2002. The work of these cooperating organizations has produced the widely used Access to Biological Collections Data (ABCD) Schema 2.06, a comprehensive standard for the access to and exchange of data about specimens and observations. The architectural body and custodian of the current recommended standards is the **Biodiversity Information Standards (TDWG) Group**, also known as the Taxonomic Databases Working Group, a not for profit scientific and educational association affiliated with the International Union of Biological Sciences

Multisatellite Thematic Mapping - This Task Group was established in 1984 to conduct a demonstration project to show how data from several satellites could be combined with direct ground observations to construct a detailed map of the surface features of an area. Three satellites with different sensor characteristics were employed. The resulting map showed the geological

features, soil types, vegetation, and other details of a region of Tanzania near the equator. The Task Group was chaired by J. P. Monget of France, and C. Bardinnet and G. Gabert participated in the mapping. The results were summarized in *CODATA Bulletin No. 62* (Oct. 1986), *Methodology for Multisatellite Thematic Mapping*. After successfully completing this prototype project, CODATA left the responsibility to the many other international groups that were becoming active in the use of satellites for mapping.

Data for Surveillance of Active Volcanoes – Marcello Carapezza, Italian Delegate to CODATA, chaired this Task Group, which was established in 1982. The group organized a *Symposium on Data for the Surveillance of Volcanic and Seismic Activities in Sicily* in 1983. Its objective was to improve the methods of correlating the measurements on gases and waters coming out of volcanic areas with the risk of volcanic and seismic activity. Unfortunately, the group became dormant after the untimely death of Carapezza in 1987.

Geothermodynamic Data – This Task Group was started in 1984 with the goal of extending the CODATA work on accurate thermodynamic properties of simple, well-defined chemical compounds to the more complex minerals encountered by geologists. Edgar Westrum was the prime mover in setting up the group, which was first chaired by I. L. Khodakovsky of the USSR, and later by S. K. Saxena of Sweden. The Task Group helped organize a large symposium in Strasbourg in 1988 on the topic of *Thermodynamics of Natural Processes*. Preliminary reports were issued but the resources were not available to produce a comprehensive database of mineral thermodynamics.

CODATA Referral Database – In 1984 CODATA started its first effort to produce an output in digital form. In a joint project with UNESCO, this Task Group converted the information that had appeared in the various print directories produced by CODATA, as well as a related directory from UNESCO, into a searchable database. The final output was a floppy diskette with the data and self-contained retrieval software. A user could load the diskette onto a personal computer and search for data sources by entering the topic, geographical location, etc. The output was a record of information on possible

sources for the data, including contact information, cost, and technical details. The first version was released to the public in January 1990. Many copies of the diskette were distributed, and some were sold. Well before the Internet and Google, this project illustrated the practicality of distributing data and associated software on digital media to individual scientists for access with their personal computers.

The Task Group was first chaired by Edgar Westrum, and later by David Watson; the main participants were F. J. Smith of Belfast University, Bettijoyce Molino of the National Bureau of Standards, and D. Laurent of the University of Paris. John Rose of UNESCO took an active part in the work.



CODATA Referral Database Task Group – Jacques Dubois, David Lide, Edgar Westrum, John Rose (at terminal), Jack Smith, Bettijoyce Molino, David Watson

Artificial Intelligence and Computer Graphics (AIGRA) – This Task Group was established in 1986 under the direction of Jacques-Émile Dubois, with a broad objective to explore new ways to derive useful knowledge from scientific databases. It conducted a workshop on *New Perspectives in Scientific Complex Data Management* in Paris in 1989; the proceedings were published as *CODATA Bulletin*, Vol. 22, No. 4, 1989. The Task Group also contributed to an Advanced Materials Workshop held in 1992 at Como, Italy, focusing on an interdisciplinary approach to materials property prediction from first principles rather than by trial and error testing – a potential saving of both time and expense in the development of new materials with specific properties. Another workshop was held in 1994 on the role of data in decision making in science and technology. The topics addressed ranged from aircraft design to environmental

protection, including the role of databases in management decisions.

Industrial Data Commission – This Commission was established in 1984 to guide the Executive Committee on the data needs of industry. It was chaired by Jack Westbrook, a US physicist who had a long career as an industrial materials expert for the General Electric Company, and included members who worked in the chemical, aeronautics, metallurgical, and other industries. Among other activities, it conducted an *International Workshop on Materials Data Systems for Engineering* in Schluchsee, Germany, in 1985 that addressed industrial needs and recommended several new CODATA activities. One of its recommendations was to set up a Task Group on Materials Database Standards, later broadened to include other aspects of database management.



Industrial Data Commission – Seated, left to right: Jack Westbrook, David Abir, Fedor Kuznetsov; Standing: Reiner Eckermann, James Naylor, Peter Schönholzer, Keith Reynard, Jiro Hiroishi

Materials Database Standards/Management – CODATA cooperated with the US National Bureau of Standards (NBS) in sponsoring an important workshop on computerized materials data in November 1982 at Fairfield Glade, Tennessee. The meeting was organized by Jack Westbrook (US) and John Rumble (US). The results were published by NBS as *Computerized Materials Data Systems: The Proceedings of a Workshop Devoted to Discussion of Problems Confronting Their Development*. As a direct result of this workshop, several programs to build online materials data systems were started. In addition, at the recommendation of the Commission on Industrial Data, the Executive Committee set up the Task Group on Materials Database Standards in 1988; it was first chaired

by Anthony Barrett, former Director of the Engineering Sciences Data Unit (ESDU) in the UK, and later by J. G. Kaufman, head of the National Materials Property Data Network in the US. One of its first actions was to start a *Materials Data Newsletter* which appeared as a part of each issue of the *CODATA Newsletter*. The newsletter provided an overview of national and international developments relating to all aspects of materials property databases. It served as an effective communications center for the growing number of materials databases being developed throughout the world, as well as for various national and international standardization bodies. Other outputs included the publications:

- *Guide to Material Property Database Management* (CODATA Bulletin No. 69, 1988),
- *The Provision of Materials Property Data via Computerized Systems* (CODATA Bulletin, Vol. 22, No. 2, 1990), and
- *International Register of Materials Database Managers* (CODATA Special Report No. 13, 1990).

Through its newsletter, publications, and sponsorship of conferences, the Task Group played an influential role in the growth of computerized materials data dissemination in the 1990s. It cooperated with ASTM, CEC, and other organizations in organizing conferences on materials data standards (e.g., in Gaithersburg, Maryland, and Frankfurt, Germany, in 1993; Tsukuba, Japan, in 1995) and promoted studies on the economic benefits of computerized dissemination of materials data.

Chemical Thermodynamics Tables – Thermodynamic data have been an important theme since the founding of CODATA. Task Groups on Key Values and Systematization of Thermodynamic Tables influenced the work on table and database development in many countries. A successor Task Group under the name Chemical Thermodynamic Tables was approved in 1982; it was chaired first by Henry Kehiaian of the University of Paris, later by Malcolm Chase of the National Bureau of Standards. The group prepared a set of prototype tables for calcium compounds and published it as a book: *CODATA Thermodynamic Tables. Selections for Some Compounds of Calcium and Related Mixtures*, D. Garvin, V. B. Parker, and H. J. White, eds., Hemisphere Press, 1987. The Task Group also conducted a very successful *CODATA Symposium on Chemical Thermodynamic and Thermophysical Properties Databases* in Paris in 1985; the papers were published as *Thermodynamic Databases* (CODATA Bulletin 58, 1985).

As the tools of information technology expanded in the next decade, CODATA's extensive work on thermodynamic tables set the stage for the International Union of Pure and Applied Chemistry to undertake a project to establish standard formats for the computer storage of experimental, predicted, and critically evaluated thermodynamic data. This led to IUPAC adoption in 2006 of ThermoML, a format based on Extensible Markup Language (XML). This format, which greatly facilitates the storage and analysis of thermodynamic and thermophysical data, is another example of CODATA starting a prototype effort which is then taken over by a discipline-based organization for full implementation.

Data Sources in Far-Eastern Countries – As a reflection of the growing participation of Asian scientists in the work of CODATA, this group was established in 1987 under the chairmanship of Jiro Osugi of Japan; the original participants came from Japan, China, Korea, Indonesia, and Thailand. Its overall goal was to expand the long-standing CODATA effort on directories of data sources to include East Asia. The group faced special problems in dealing with language barriers and less familiar publication outlets, but was able to issue its first directory in 1989 as *CODATA Bulletin* Vol. 21, No. 3. This



Task Group on Data Sources in Asian-Oceanic Countries

listed 90 Chinese, 117 Japanese, and 9 Korean data sources, and it was followed by a large number of other publications in the next few years. In 1990 the group was

renamed Task Group on Data Sources in Asian-Oceanic Countries (DSAO) and its coverage expanded to The Philippines, India, and Pakistan. A consolidated volume appeared in 1994: *The CODATA Directory of Data Sources in Science and Technology in Asian-Oceanic Countries*, Yaruo Hu and Edgar Westrum, eds., CODATA Monograph Series, Vol. 2.

After the death of Jiro Osugi, Mitsuo Tasumi became Chairman of the group. Other active participants were Yaruo Hu of China-Beijing, Jen-Leih Wu of China-Taipei, Makoto Kizawa of Japan, and Edgar Westrum, the long-standing leader of the CODATA publication program. In addition to providing means of access to data sources that were largely unknown in western countries, the work of this Task Group brought many new scientists from Asian countries into the CODATA family.

Publications and Other Outputs

A number of publications resulting from Task Group activities in the 1980s have already been described, some as issues of the *CODATA Bulletin*, others as separate books. Proceedings volumes from each of the CODATA Conferences were also published; in some cases groups of closely related papers from a Conference were published as independent volumes. Two very important

publications came from Task Groups established earlier and discussed under the previous decade:

- release of the updated set of Fundamental Constants in 1986; and
- final publication of the *CODATA Key Values for Thermodynamics* in 1989.

Finally, the following books prepared under CODATA sponsorship should be noted:

- *Database Management in Science and Technology*, John Rumble and Viktor Hampel, eds., North-Holland, 1984. This book described the technical details of planning, designing, and implementing databases of numerical scientific data. It became a valuable sourcebook for groups beginning the transition from paper records to computer management.
- *Nonbibliographic Data Banks in Science and Technology*, Stephan Schwarz, David G. Watson, and Olov Alfeldt, eds., ICSU Press, 1985. This book resulted from a training course conducted by CODATA, UNESCO, and the Swedish Delegation for Scientific and Technical Information.
- *Directions for Internationally Compatible Environmental Data*, G. C. Carter and Barry I. Diamondstone, eds., Hemisphere Press, 1990. This book, based on a CODATA-sponsored workshop, addressed the measurement methods and standards needed to assure environmental monitoring data from throughout the world can be compared with confidence. Selected papers from the workshop appeared in the *CODATA Bulletin*, Vol. 21, No. 1-2.
- *CODATA Guides for Data Publication*, Edgar Westrum, ed., *CODATA Bulletin* Vol. 21, No. 4, 1989. This was a consolidation and update of guides to the presentation of data in the primary literature that were previously published individually by CODATA.

Summary

The decade of the 1980s was an eventful one for CODATA. It marked a transition from the initial CODATA focus on the physical sciences to a more balanced program involving all the natural scientific disciplines as well as relevant areas of engineering. This transition is reflected in the list of scientific activities given above, as well as in the active participation of bioscientists such as Alain Bussard, Boja Keil, Harald Haendler, Lois Blaine, and Micah Krichevsky, and geoscientists such as William Hutchison, Richard Sinding-Larsen, Marcello Carapezza, and James Crease. Many of these served as Officers or Executive Committee members.

The record of the decade also shows that CODATA responded rapidly to the dramatic advances in the use of computers for data management and dissemination. Through its Conferences, workshops, and publications it played a significant role in educating scientists on the new tools available through modern information technology. It started operational projects like the Hybridoma Data Bank, Microbial Strain Data Network, and CODATA Referral Database. This set the stage for the advance to the Internet that took place in the 1990s.

Another trend was the increasing collaboration of CODATA with other organizations concerned with scientific and technical data. These joint activities involved many ICSU Unions and other ICSU bodies; government agencies and private bodies in member nations; and, especially, UNESCO and other United Nations bodies. The work of the CODATA Commission on Standardized Terminology for Access to Biological Data Banks was an outstanding example of the impact of CODATA in bringing together multiple organizations to solve a problem that was important to all of them. Through these collaborations CODATA was able to leverage its limited financial resources to achieve far more than it could have done alone.

Finally, the decade saw CODATA membership and, more generally, active participation in CODATA projects expand from its initial base of Europe and North America (plus Japan) into all of Asia. The admission of the two Chinese Academies, to be followed later by Korea and Indonesia, brought new scientific talent and enthusiasm into CODATA.

The 1991-2000 Decade

Governance

The 18th General Assembly convened in Beijing in October 1992 with 28 National and Union Members represented. This gathering represented two significant firsts – the first time the General Assembly had met in China and the first time that the Chinese Academy in Taipei was represented, having been accepted as a National Member at the conclusion of the 17th General Assembly in 1990. It is of interest to note that the application of Russia to assume the National Membership of the former USSR was approved and the Union of Biochemistry and Molecular Biology was approved as a new Scientific Union Member. James Crease and Lev V. Gurvich were re-elected as Treasurer and a Vice President, respectively, and Akira Tsugita (Japan) was elected as a Vice President. A renowned researcher in bioscience, Tsugita had helped lead the way for CODATA to engage the life sciences. One of his earliest



Akira Tsugita

contributions to CODATA was a leading role in the Task Group on Protein Sequence Databases which forged a means of collaboration among the three major international protein sequence research groups which hitherto had been working independently.

An important Long Range Planning meeting preceded the 19th General Assembly in 1994 at Chambéry in southern France. As it had been some ten years since the previous planning meeting, it was thought to be time to re-assess CODATA's programs,



Strategic Planning Meeting – Chambéry France 1994 – David Lide presiding

vigour, and future options. The CODATA Executive Committee and several invited experts spent two days developing a plan to guide CODATA's activities for the next decade.

The following is an extract of the ensuing report, prepared by David Lide, who chaired the planning meeting.

Although the challenges and opportunities of the 1990's were very different from those faced by CODATA at its inception in the 1960's there was still a need for an organization with broad international and interdisciplinary representation and a mandate to address issues of data management and dissemination across the full spectrum of science and technology. Consequently, in order to capitalize on its unique structure and mandate, CODATA needed to broaden its sights and give priority to activities that affect all scientific disciplines and will have maximum impact on the scientific community. There was a concomitant need to expand its collaboration with the ICSU Unions and other scientific bodies, seeking opportunities for symbiosis between the data management expertise within CODATA and the subject matter expertise in more specialized organizations. Less emphasis should be given to narrow, discipline-specific projects that could equally well be carried out by the Unions alone; the goal should be to collaborate with the Unions in cases where CODATA can make a unique contribution.

In order to attain the goal of becoming the central focus within the ICSU family and the scientific community at large for the methodology of organizing, managing, assuring quality, and disseminating data from all scientific and technological disciplines, CODATA needed to emphasize the following strategic elements:

- *Electronic Dissemination of Data – greater use of the Internet and other electronic networks to disseminate*

CODATA's outputs, such as key data sets, directories and guidelines, as well as assisting the scientific community to adapt to the transition from print to electronic data dissemination;

- *Studies of Abstract Data Structure – fundamental theoretical studies of the generic characteristics of data and translation of the results into forms suitable for application to individual scientific disciplines;*
- *Data Activities in Systematic Biology – collaboration with the IUBS and other biological organizations to develop an integrated series of databases which catalogue the known species on the earth;*
- *Access to Environmental Data – cooperation with the World Data Center Panel and other ICSU bodies to solve problems of access to geophysical and environmental data pertinent to global change research;*
- *Enhanced Outreach Activities – expansion of educational and tutorial activities aimed at helping the scientific community take advantage of modern information technology.*

The basic organizational structure of CODATA was seen to be sound but the following operational changes were recommended to make CODATA function more efficiently and effectively.

- *The Executive Committee should make greater use of sub-committees to address issues such as finding new funding resources and developing new projects as well as to investigate specific questions that arise. Use*

should be made of the authority to co-opt members to add disciplinary expertise when needed.

- *Task Group proposals should be examined more critically and Task Group performance should be monitored carefully to assure that CODATA resources are being used optimally. It may be desirable to reduce the number of Task Groups and provide more support to the most effective ones.*
- *A serious effort was needed to obtain more publicity for CODATA by, for example, periodic press releases and the use of the Internet to publicize CODATA activities – requiring of course, tangible outputs from successful CODATA projects.*
- *Systematic steps needed to be taken to find new financial and human resources for CODATA by establishing relationships with other bodies.*

The 28 National and Union Delegates to the General Assembly elected Jacques-Émile Dubois (France) as President and re-elected Gordon Wood as Secretary General, both by acclamation. Reflecting CODATA's interests in the quickly developing World Wide Web was the decision to form two Working Groups: one to examine issues related to international access and exchange of data for scientific research on the Web; the other to assist in the immediate development of a CODATA Web site in Canada.

Tsukuba, Japan's 'Science City', was the site of the 20th General Assembly in October 1996 with the customary number of Delegates present. Notable among the actions and decisions were the admission of Indonesia and Senegal as National Members and the election of Michael

Chinnery (US) as Treasurer and Fedor Kuznetsov (Russia) as Vice President. The Task Groups approved at this General Assembly reflected CODATA's ongoing concerns and emphases related to the impact of the computer on data accessibility and exploitation: Data/Information Visualization; Outreach and Education; the use of the World Wide Web (Internet); and Access to Data.



Michael Chinnery

An expert in geo-data, Michael Chinnery, had recently retired as Director of the US National Geophysical Data Center in Boulder, Colorado having served as Director since 1982. He was also Treasurer of the ICSU Panel on World Data Centres and involved with the Earth Data Network for Education and Scientific Exchange (EDNES).



Fedor Kuznetsov

Kuznetsov, Director of the Institute of Inorganic Chemistry in Novosibirsk, was leading a strong program in several areas of materials science: theoretical modeling; searching for new micro-electronic materials; developing quantitative methods for analyzing inorganic materials; and pursuing novel preparation techniques. A prominent physical chemist and materials scientist, he was instrumental in the development of a data bank on properties of electronic technology materials.

The 21st General Assembly convened in New Delhi, India, in November 1998 with 25 National and Union Delegates attending. John Rumble, Jr. (US) and Paul Mezey (Canada) were elected as President and Secretary General respectively. Akira Tsugita was re-elected as Vice President.



John Rumble instrumental in organizing numerous data-related workshops on behalf of CODATA. At the US National Institute of Standards and Technology (NIST),

he led development of scientific and technical database standards, including the major standard for industrial data exchange developed by the International Organization for Standardization (ISO). Under his leadership NIST produced and sold more than 70 personal computer compatible databases of scientific information. He effectively guided NIST into the Internet era with the release of more than 25 online data systems in virtually every area of science and engineering.



Paul Mezey

Representing the perspective of a data user more than a producer, Mezey's expertise involved creative applications of data in his field of biomolecular modelling and simulation which provide powerful tools for molecule design, molecular engineering, and molecular level biotechnology.

Typically, the approved Task Groups mirrored the current interests and concerns of extracting knowledge from data collections, maintaining quality, ensuring database compatibility, and promoting standardization of formats. Of particular interest was the formation, in cooperation with ICSU, of a Group on Data and Information reflecting ICSU's interest in receiving advice regarding data issues across the whole spectrum of its disciplines.

As one means of raising the profile of data science it was decided to initiate a biennial CODATA Prize, to be awarded at future General Conferences to individual scientists nominated by their colleagues in recognition of exceptional contributions in the area of scientific data. Another indicator of the impact of the Internet on the world of data was the agreement to explore the implementation of an electronic CODATA Data Journal.

In 1999, the CODATA Officers and several other individuals met in Chantilly, France, to discuss how CODATA could meet several challenges, including a change in the Executive Director position (see below), the emergence of the next generation Task Groups, and impending financial shortfalls. While this meeting did not produce a new strategic plan, a consensus was reached about facing these operational challenges.

One result was the decision to begin holding Executive Committee meetings in Member countries every second year with a view to creating stronger ties with National Members. A second result was the recognition that the International Conferences needed more organizational oversight by the Secretariat and that International Program Committees would be established to set the Conference agendas.



Jean-Jacques Royer

was more a user than producer or evaluator of data, thus bringing balance to the Executive Committee. Nigeria was welcomed as a new National Member and it is of interest to note that Ukraine and Thailand were accepted as National Members via a Postal Ballot following the Conference. Eight minor amendments to the Constitution were approved, the most significant one being an update to the so-called 'Postal Ballots' to include telefacsimile and email as legitimate means of voting.

Secretariat



**Phyllis
Glaeser**

A very significant change at the Secretariat was the retirement of Phyllis Glaeser from her position as Executive Director in June 1999 after having been with CODATA for a quarter century. Blessed with an instinctive ability to help scientists from many nationalities and diverse disciplines to work together harmoniously, she served as the hub around which CODATA's various spokes turned. She was comfortable with the rapid changes occurring in informatics and served as a mediator between partisans of the classical ways of doing data science and those impatient to adopt all the newest techniques and tools.

During her tenure she was responsible for numerous editions of the CODATA Handbook along with many Conference Proceedings and *Newsletters*. Indeed, Phyllis had stamped her personality on CODATA in its formative years, integrating people, disciplines, and pioneering techniques.



Kathleen Cass

Thanks to proactive recruiting on the part of then-President John Rumble, Kathleen Cass was appointed Executive Director in July 1999. An Irish lawyer, having practised in both Dublin and Paris prior to working with the International Union of Biological Sciences for two and one-half years, Kathleen brought an impressive skill set to the position. With many details unresolved regarding the organization of the CODATA Conference in 2000, Kathleen barely had time to remove her coat before plunging into work on the Conference and associated General Assembly.



**Edgar
Westrum**

Another very significant change in 1999 was the retirement of Edgar J. Westrum as Editor-in-Chief of CODATA. Working tirelessly in this volunteer position, Edgar was associated for many years with the *CODATA Bulletin* series as well as the *Newsletter*, in addition to serving as Secretary General for 8 years (1974-1982). With his ever-present camera, Edgar was renowned for capturing the highlights of and participants at countless CODATA events. In the day before it was trivial to send a photo via email, many a CODATA colleague received a photo print of a recent event in the post from Edgar.

CODATA Conferences

The 13th Conference in Beijing, October 1992, entitled *New Data Challenges in Our Information Age*, was especially notable having attracted 283 participants, with approximately 50% of those coming from 21 countries

other than China, one of the largest attendances to date of a CODATA Conference. Several joint sessions with the ICSU Panel on World Data Centers, which met concurrently with CODATA, were a novel feature. Presentations included five plenary lectures, about 80 lectures in 21 discipline-oriented sessions, and a combination of about 120 oral discussions and posters in a series of six topical workshops. Some idea of the scope of topics covered is conveyed in a partial list: data handling; social, political and legal aspects of data; materials databases; genome projects; molecular modelling; emerging methods for data handling; artificial intelligence; and data exchange formats.

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. 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 will be the integration of existing and future data sets because the detection of change is virtually impossible if data sets gathered from a variety of sources cannot be accurately registered geographically.

An afternoon tour of the famed Peking Man Site at nearby Zhoukoudian, preceded by a plenary lecture given by Wu Rukang of the Chinese Academy of Sciences on current excavation and research activities, was well received by many.

Data and Knowledge in a Changing World: the Quest for a Healthier Environment was the title of the 14th CODATA



Conference Delegates – Chambéry, France 1994

Conference at Chambéry, France, in 1994 with some 292 attendees. The Conference had two principal objectives: to update research and the choice of essential data in two associated fields – innovation (search for new materials) and environment (biodiversity and toxicity) – and to highlight developments in computer-aided systems and communication as they relate to exploiting and sharing data. A sample of session titles gives some indication of the scope of the conference: “Computer-aided design of composite materials”, “Biodiversity: taxonomic and information structures”, “Obstacles to open exchange of scientific data”, “Information requirements in developing countries”, and “Environmental issues: national and international policies”.

The 15th CODATA Conference, *Scientific Data in the Age of Networking*, was held in Tsukuba, October 1996 with over 200 participants, representing more than 20 countries. In his keynote address, Y. Nishijima, President of the Science Council of Japan, reminded the audience that in many instances it is the integration of knowledge from various fields that has led to significant scientific and technical advances – integration facilitated by the networking age. Ferris Webster, Chair of CODATA's newly formed Commission on Access to Data, suggested possible roles for CODATA in dealing with the current flood of data and information relating to geophysics and the environment. The three Kotani Memorial lectures, named in honour of Masao Kotani, one of CODATA's founders and a past CODATA President, were another highlight.

The variety and scope of papers given in discipline-oriented sessions reflected the multi-disciplinary nature of CODATA. Life Sciences sessions included a focus on genome research, biological macromolecules (featuring reports on compilations of protein sequences and of protein and nucleic acid structures) and species diversity (emphasizing the harmonization of informatics and biology). The Physical Sciences stream covered topics ranging from space to nanostructures with the reminder that the Internet, while providing a powerful mechanism for the direct dissemination of data from the collector to user, raises many new problems related to the quality of both data and metadata. A fascinating example of the progress being made in the practical application of materials data and computer modelling was a presentation

from Boeing Information and Support Services stating that their new model 777 aircraft was the first they had manufactured for which no physical mock-ups had been required.

Convened November 1998 in New Delhi, the 16th CODATA Conference focussed on *Scientific and Technical Data and Communication for the Sustainable Development of Nations – Data Management in the Evolving Information Society*. Approximately 170 participants, 70 of whom represented 20 countries along with 100 from India, attended the 13 plenary oral sessions and two poster sessions. The exchange of information and ideas covered recent revolutionary developments and applications reflecting the innovative changes and rapidly increasing importance of data science, including databases, management, training, and sources in the Computer Age. In this context the valuable role played by CODATA in fostering conscientious use and development of Data Science in an international setting was evident – a practical demonstration and reinforcement of CODATA's *raison d'être*.

Nature intervened in a major manner with regard to the actual site of the 17th CODATA Conference in October 2000. Planned for Baveno, Italy, the Conference began there and continued for two and a half days until torrential rains caused nearby Lago Maggiore to flood the host hotel, requiring it to be evacuated. Thanks to heroic efforts on the part of the hosts, the approximately 200 conference attendees were relocated to a site in nearby Stresa within three hours where the Conference continued without missing a session and the General Assembly was held without further incident.

Reflecting the concerns of the time, this Conference focussed on the need for new insights into the collection, supply, management, and analysis of data in all fields ranging from biotechnologies to computer-aided medical assistance to the immense problems related to global climate change. Attention needed to be paid to aggregating the vast amount of data being generated and disseminated, and making those data interoperable among a plethora of systems coming online. More specifically, the conference focussed on issues such as: understanding the data and information elements implicit in current scientific activities; identifying the challenges

associated with managing the vast amount of data being generated; recognizing the data needs of the 'soft' sciences such as sociology and economics; and ascertaining the quality and quantity of data needed by planning agencies in the fields of science, technology, and medicine.

A highlight of the 17th Conference was the awarding of the first CODATA Prize to Barry Taylor of the US National Institute of Standards and Technology in recognition of "his major contributions to the advancement of our understanding of the physical world through critically evaluated values of the fundamental constants". The painstaking, detailed adjustment of the value of these constants provides a self-consistent set for use by scientists and engineers throughout the world in essentially every field. The work of the CODATA Task Group on Fundamental Constants, which Dr. Taylor chaired for 6 years was, and continues to be, a 'flagship' effort providing high worldwide visibility for CODATA.



Barry Taylor

Scientific and Technical Activities

New Task Groups approved in 1992-2000

Databases for Experimental Data & Electronic Publishing (EXPERIDAT) – The aim of this Task Group, approved in 1992 and proposed and led by Ekkehard Fluck, head of the Gmelin Institute for Inorganic Chemistry, was to establish standards for formatting experimental data in the physical sciences and engineering in order to facilitate automated or semi-automated input into electronic databases as well as the transfer of data among such databases. The Task Group aimed to address an important issue facing scientists, publishers of scientific journals, and producers of scientific numeric databases as they sought to exploit the rapidly expanding advantages of the computer in evaluating, organizing, and disseminating factual data. More specifically, the Group proposed to explore the potential to reduce size, and therefore costs, of paper publications, to reduce costs of producing electronic databases, and to speed up availability of new data to the

scientific community. In a sense, the goal was to help the physical science community catch up to the molecular biologists who had been submitting sequence data electronically to their journals and databases for some time. It turned out that, by 1996, the efforts of this Group were subsumed by a group set up by ICSU supported by offers of expertise from the International Astronomical Union.

Property Data for Safety, Environment and Design (1992) and Distributed Data Depository for Experimental Thermophysical Property Data (1994)

– The focus of this re-named Task Group, a metamorphosis of the former Task Group on Critically Evaluated Phase Equilibrium Data of Organic Substances, was to facilitate international cooperation in the establishment of a CODATA/IUPAC World Data Depository of published and unpublished experimental thermophysical property data for compounds and mixtures, avoiding the duplication of collecting public data. Kenneth Marsh, of the Thermodynamics Research Center in the US and leader of the Group, noted two significant initial aspects to be addressed: (1) the establishment of an international, easily accessible, non-commercial, electronic, archived database wherein scientists could deposit their data at the publication stage; (2) the establishment of a resource providing the many national bodies involved in data evaluation with ready access to greater amounts of scientific data than would be available if they were to consult only national databases. The Group conducted a workshop on *Thermophysical Property Data Banks and Process Simulators* in 1994.

Data Access Commission – Established initially as a Working Group by the Executive Committee, this activity was approved in 1994 as a Commission because of its broad, cross-cutting range. Chaired by Ferris Webster, an oceanographer at the University of Delaware with a wealth of experience in international collaborative endeavours, this Commission was charged with advising CODATA and ICSU on the issues related to ready access to numerical data at a time when storm clouds of restrictions arising from financial, political, and geographic pressures were threatening. While in Working

Group status, they published their report *Data Access*, number 15 in the CODATA Special Report series, and prepared a draft ICSU brief on data access for submission to the World Meteorological Organization (WMO) Congress in 1995. This Commission input helped to clarify WMO policy on unrestricted access to meteorological data by the research and educational communities.

To reflect more clearly its broad base and the integration of science, technology, and legal issues, the Commission joined forces with the ICSU ad hoc Group on Data and Information and became known as the ICSU/CODATA Group on Data and Information in 1997. This Group addressed the efforts that had started at both national and international levels to impose legal restrictions on access to data, which traditionally were not subject to protection by the usual copyright laws. They participated actively at WIPO (World Intellectual Property Organization) meetings in 1997 and 1999, presenting an important paper on database property rights giving emphasis to the needs of the scientific community. They also monitored efforts in the United States and the European Community to impose restrictive protection on scientific and technical data. In October 2000, the Group conducted a follow-up workshop in Baveno to review what had happened since the EU Database Directive took effect in January 1998 and to prepare to make substantive input to the EU review process. The Group served a very important service by alerting the scientific community to the danger imposed by attempts to place legal restrictions on the free exchange of scientific data. Without the efforts of this group, it is entirely possible that traditional scientific practices on data exchange would have suffered.

Global Plant Checklist Network – That the proponents of this Group came to CODATA in 1994 for backing was an indication of the credibility CODATA had developed as an umbrella organization serving as a catalyst in bringing groups together and initiating large international, inter-Union projects. Task Group Chair Karen Wilson, from the Royal Botanic Gardens in Sydney, Australia, noted that the Group would involve botanists and computer experts around the world in order to reach their goal of looking at all plants, of which there were

estimated to be some 300,000 species. The Network would be relevant to industry and serve as well as a link for virus databases and a nomenclature source for molecular databases. By establishing this Group, it was clear that CODATA was moving further into the biological field and broadening its reach well beyond the original fields of chemistry and physics.

Their meeting in Madrid was particularly useful as it facilitated liaison between members of the Task Group and the Species 2000 Steering Committee as well as the International Working Group on Taxonomic Databases. During the biennium, the Task Group made informal approaches to 22 prime botanical institutions worldwide to solicit their participation as network nodes.

Continuing its collaboration with other groups, the Task Group met jointly in 1997 with the Species Plantarum Project in June and with the IUBS Taxonomic Database Working Group and the Species 2000 management team in November. Another meeting was held in October 1998 in conjunction with the International Organization for Plant Information (IOPI). Significant progress was made in dealing with such practical matters as names that had been misapplied, or cases where a single species had been split into two or more taxa, so that correct assignments could be made in the Checklist. By 2000 the Checklist contained six regional datasets (Australia, Europe, Russia, Mediterranean region, N. America north of Mexico, and Peru) and three complete family datasets (Casuarinaceae, Irvingiaceae, and Magnoliaceae).

Thermodynamic Data for Key Chemical Substances

– A successor to the Key Values for Thermodynamics Task Group which completed its work in 1980, having established the data for 36 elements, this Task Group aimed at preparing recommended data sets for another ten key elements. Approved in 1994 and led by Lev Gurvich, a proven expert in the field from the Institute for High Temperatures in Moscow, the Task Group's challenge was to distill the best and most reliable data from the plethora of obsolete or poorly-cited data of questionable provenance and quality known to exist in books and journals worldwide. The Group prepared its initial reports on Barium and Strontium, and their

compounds, in preparation for producing recommended values in 1996. However, the death of Lev Gurvich in 1995, along with the evaporation of financial support for thermodynamics work both in Russia and the US, combined to effectively nullify the activity of this Group.

CODATA and the World Wide Web – Leading the way for CODATA onto the then-rapidly expanding World Wide Web (Internet) stage, the Working Group on Electronic Information Transfer established an international site in 1994, at the National Research Council in Canada, which was designed not only as a working tool for those involved in CODATA but as a means of introducing CODATA to the world. Following the expected developmental path, this Working Group was elevated to Task Group status in 1996 under the leadership of James Crease. Three objectives were in view: to make CODATA better known internally and externally by providing information about it and by disseminating its reports and products via the Internet; to make CODATA more useful as a source of scientific and technical data and metadata by listing existing repositories and/or providing a gateway to them; and to advise Member countries wishing to establish their own home pages to be linked from the main site. By 1996, four National, four Union and three Co-opted Members, along with two Task Groups, were using the Internet to promote their activities. With its work effectively completed by 2000, there was no need for the Group to be renewed.

Data/Information and Visualization – As computer power and international connectivity were increasing at a prodigious rate the 1996 birth of this Task Group was almost inevitable. With Nahum Gershon, MITRE Corporation (US) as Chair, the Task Group aimed to be a resource to the CODATA family for exploiting the Internet by providing, for example, glossaries and applications relating to new technologies, workshops on distributed information systems, and tools for visualizing the information latent in various collections of data. During its tenure, the Group convened a number of workshops, examples being those in Paris (*Information Visualization: Where We Are and Where Do We Go from Here?*) in 1997, New Delhi (*Information Visualization and*

Management of Heterogeneous Systems) in 1998, and Ottawa (*Information Visualization Tutorial*) in 1999.

It is noteworthy that, with the approval of these two Internet- and computer power-related Task Groups, CODATA was expanding its traditional focus on data quality and evaluation to include exploiting the knowledge and information contained in those data.

Outreach, Education and Communication – Approved in 1996, this Task Group was, in essence, a response to an objective identified in the 1994 CODATA Strategic Plan, that of assuming a more proactive role in transferring the expertise of CODATA members to the scientific and technical community at large. Chaired by Micah Krichevsky, a biologist with wide-ranging interests in biological information and founder of Bionomics International, US, the Task Group planned to assess the needs in various disciplines and jurisdictions, proceeding as financial and human resources permitted. Sustainability of the enterprise called for an emphasis on building local capacities by teaching the teachers. They received funds from the ICSU Grants program for constructing a registry of scientists and technologists in all disciplines worldwide, but emphasizing developing countries. Such a registry would facilitate scientist-scientist contact – a vital step in improving collaboration and cooperation in developing countries. Liaison was established with the ICSU Committee on Capacity Building in Science and significant indications of support were received from Egypt, South Africa, and the US Agency for International Development. An article “ASM launches Online Community of Microbiologists” was published in the ASM (*American Society of Microbiologists*) *News* in 1999. Meetings were organized in Paris and Bangkok in January 2000.

Comparative Mathematical Methodologies of Data Handling and Knowledge Extraction – Building upon collaborative work initiated by several French, American, and Russian scientists, this Task Group had as its goal the development of mathematical techniques to compress and characterize large data sets while maintaining their information content and the ability to extract knowledge

or information specific to a particular application. As observed in 1998 by its founding Chair, Alexei Gvishiani, an expert in this field from the Center of Geophysical Data Studies and Telematics Applications, Russian Academy of Sciences, the work was consonant with the emerging emphases in CODATA of using evolving computer tools to mine data collections rather than simply to certify, store, and disseminate them. In 1999, the Group organized a workshop in Paris, “Mathematical and Telematics Techniques for Large Data Sets Characterization and Compression”, and published a book *Dynamic System and Dynamic Classification Problems in Geophysical Applications* (J-É Dubois & A. Gvishiani).

Standardization of Physico-Chemical Property Electronic Data Files – With decentralized publishing of numeric data on computer networks now becoming a reality, this Group, approved in 1998, sought to address the need for standard computer-readable formats for numeric data and their metadata. Such standards were essential to the scientific community as a means of documenting the network-published electronic numeric data clearly, uniformly, and completely. Working in cooperation with IUPAC and ICSTI as an influential coalition, the goal was to develop, in cooperation with all interested parties, a computer-readable standard electronic file format for data files to be disseminated via networks. Henry Kehiaian of the University of Paris VII, with extensive, pioneering experience in this field, acted as project coordinator. That this project was awarded \$100,000 in May 1999 from the hotly contested ICSU Grants Program was an indication of its anticipated value to the community.

Data Quality and Database Compatibility: Development of General CODATA Safeguards – Indicative of the concerns relating to the increasing use of computers in producing and disseminating scientific data, this Task Group sought to assist CODATA in identifying safeguards to facilitate transferability of data while simultaneously certifying quality and ensuring compatibility of data generated across different disciplines. Approved in 1998, and chaired by Paul Mezey, they organized a symposium on *Diagnostic Methods for*

Data Quality and Database Compatibility in 1999 in Budapest, Hungary. Task Group members presented 12 lectures at various conferences and symposia on data quality problems and published two scientific papers on methods of error detection in molecular shape databases. Their own meetings involved data quality and database management experts from the fields of pharmaceutical bioinformatics and related industries. Perhaps because its objectives exceeded the human and financial resources available to it, the Group did not continue past 2000.

Reliable Scientific Data Sources in Africa – By establishing this Group in 2000 – the first multi-national project within Africa to address scientific data issues - CODATA was strategically extending its reach to a continent of emerging scientific activities. One of the Group's initial goals, under the chairmanship of Abdoulaye Gaye from Senegal, was to identify and evaluate, from the perspective of quality and usability, all existing data and information sources that might be used by government, academic, and private sectors. Two workshops were held in Dakar, Senegal. Building on those findings, the Group planned to identify areas where data and information sources were lacking, required improvement or needed assistance in being disseminated.

Data on Natural Gas Hydrates – The development of a comprehensive information system on all aspects of natural gas hydrates – natural gas and water frozen together into a solid substance, common in permafrost and deep ocean sediments – was the objective of this Group. The impetus for such a system stems from the fact that the resource value of these hydrates is generally estimated as greater than the known reserves of oil, natural gas, and coal in the world. Chaired by Fedor Kuznetsov, upon its initial approval in 2000, this multinational Group aimed at finding ways to use modern information technologies for data and knowledge sharing to: achieve understanding of the role of natural gas hydrates in natural processes; find rational ways of exploiting gas hydrates resources; and prevent hazards resulting from uncontrolled eruptions of gases from gas hydrates deposits. Having received an ICSU grant as a measure of the importance of their work, the Group developed and published the Gas Hydrate Markup Language (GHML) in 2006 with a view to facilitating the

interchange of rapidly increasing amounts of gas hydrate data among researchers worldwide. Their ultimate goal was the establishment of the Gas Hydrate Portal allowing any nation, with a minimum of effort, to add their database to the international network but that had not yet been accomplished by 2010.

Global Species Data Network – Essentially a unification in 2000 of the pre-existing Standardized Terminology for Access to Biological Data and the Global Plant Checklist Task Groups, this Task Group had the ambitious goal of developing a uniform and validated quality index of names of all known species for use as a practical tool. According to co-chairs Frank Bisby, University of Reading, UK, and Karen Wilson, the names of organisms are the key to biodiversity communications and as such, provide access to the accumulated knowledge of all life on Earth. However, despite the obvious value of a catalogue, no comprehensive indexing system yet existed for the 1.75 million animals, plants, fungi, and microorganisms named by science. This lack of a widely accessible index, with built-in mechanisms for maintenance and updating, was seen as a significant constraint on all nations wishing to fulfill their obligations under the Convention on Biological Diversity. It should be noted that, because of CODATA's seed funding, this Task Group, in cooperation with the Species 2000 group described in more detail below, was able to tap into much larger funding sources thus illustrating another important aspect of CODATA's contribution to international science.

Fundamental Constants – In 1998, the Task Group released its latest adjustment to the values of the basic constants and conversion factors of physics and chemistry, the second such adjustment since the initial publication in 1973. Complementing the traditional means of publication, the recommended values were posted on the Internet and soon thereafter consulted by close to 50,000 users – a tangible indicator of the value of the work of CODATA's renowned Task Group. With a view to succession planning, NIST engaged Peter Mohr, a distinguished theoretical physicist, to work on the constants along with Barry Taylor.

Other Developments of Interest

As another example of CODATA serving as an effective umbrella organization, the International Union of Biological Sciences (IUBS) and the International Union of Microbiological Societies (IUMS) recognized the need for an organized approach to dissemination of a synonymic checklist covering all the known species on earth. The vision was to link the myriad sources of taxonomic data distributed throughout the world for access via a central portal. CODATA was approached in 1991 and agreed to cooperate with this project as it fit well within its mission. The initial workshop to assess the project was held in Manila, Philippines at the headquarters of FishBase. The workshop was funded by the United Nations Environment Program and the Global Environment Fund. The project, now registered as a “Company Limited by Guarantee” in England, was named Species 2000 with the aim of cataloguing the names and synonyms of the 1.75 million species of the world. A major challenge for cataloguing species names is the large number of synonyms that may be applied to a taxonomic entity. Application of existing synonyms requires a high degree of data curation by experts in each field.

The plan was to design a distributed architecture to link individual taxonomic datasets maintained by their curators throughout the world. Prescribed standards for content and architecture enable interoperability of the datasets that comprise Species 2000. Both an online database and an annually updated CD-ROM were the vehicles of distribution. Membership in Species 2000 was open to any organization that collects and curates taxonomic data.

The Species 2000 organization formed a partnership with the Integrated Taxonomic Information System (ITIS) housed at the Smithsonian Institute in Washington, DC. The joint global product, called Catalog of Life, was organized at centers in Reading, UK; Washington, DC; Amsterdam, the Netherlands, and Los Baños, Philippines. Regional nodes have been formed in North America, Europe, China, Australia, New Zealand, South Africa, Brazil, and Colombia. The project has an open community model in which the custodians of participating databases are also partners in the governance and

ownership of the system. Major natural history museums throughout the world as well as networked taxonomic databases and private custodians of databases continue to be active participants in the organization.

The Species Checklist is now composed of more than 1.1 million species, plus 86,000 infraspecific taxa, and lists 720,040 synonyms (2,025,196 scientific names of all ranks). Close to 100,000 new names are added to the database each year with the goal of covering 100% of known species by 2013. The Species 2000/Catalog of Life is indeed one of the most successful projects undertaken by CODATA in the biological sciences. It has been a vibrant and valuable tool for research scientists for 20 years and prospects for its long term survival are excellent.

Information Technology Activities

A 1996 survey on issues related to the trans-border flow of scientific data, conducted by the US National Committee for CODATA, reflected the concerns of the period regarding the effective use of the Internet. In essence, the survey indicated that the scientific community then had the technological power to accumulate and disseminate vast amounts of data quickly and cheaply but lacked the social and economic infrastructure to exploit that power as fully as possible. A follow-up conference dealt with the exchange of scientific and technical data among different computing environments and across diverse scientific and engineering disciplines – issues which were regarded as presenting major hindrances to the full exploitation of computer-based modeling, the Internet, modern scientific databases, and new computer technology.

Another topic which came to the fore was concern relating to the database protection implications of forthcoming European Union Database Directives. If implemented as they appeared in draft form, the Directives had the potential to limit severely the benefits of the ‘new age’ of data and information sharing made feasible by the Internet and increasing computer power. Hence, CODATA was quick to collaborate with ICSU in lobbying against the potential deleterious effects of the Directives.

That information visualization was becoming an increasingly important research area is reflected in the reports of the CODATA Euro-American Workshop on Data and Information Visualization held in Paris in 1997. The growing amounts of data readily available on fast digital networks, combined with the development of faster and larger computers, created an escalating demand for querying, accessing, retrieving, and utilizing those data to derive new information. Tools were being developed for analyzing large quantities of data, teasing out relationships undetectable to the unaided human mind, and presenting them in visual form using powerful graphical and imaging techniques.

Publications and Other Outputs

The decision in 1994 to replace the *Bulletin* with a series of monographs was a very significant change in CODATA's publishing practice. In spite of its long, successful history, the *Bulletin* was no longer attracting a reasonable subscription market and it was thought that the community would be better served by monographs which could be marketed individually according to their subject. Similarly, Conference Proceedings, though acknowledged as important channels for the scientific output of CODATA, were presenting a publishing dilemma because of the broad, multi-disciplinary range of topics at CODATA Conferences. Hence the decision was taken to discontinue publishing general Conference proceedings and instead, to select papers from Conferences, along with papers especially commissioned, and market them in a series of focussed monographs entitled *Data and Knowledge in a Changing World*.

The first three books in the series, inspired by papers at the 1994 CODATA Conference in Chambéry, were edited by J-É Dubois and N. Gershon and appeared in 1996 with the titles: *The Information Revolution: Impact on Science and Technology*; *Modelling Complex Data for Creating Information*; and *Industrial Information and Design Issues*. By March 1999, three more volumes in the series had been released: *Geosciences and Water Resources*, edited by C. Bardinnet and J-J Royer; *Thermodynamic Modelling and Materials Data Engineering*, edited by J.-P. Caliste, A. Truyol and J. H. Westbrook; and *Dynamic Systems and Dynamic Classification Problems in Geophysical*

Applications, edited by J-É. Dubois and A. Gvishiani.

The need for printed data sources directories, which suffered from the ongoing burden of being kept current, gradually began to wane as Internet-based search engines facilitated finding data of interest worldwide. With greater accessibility and wider distribution, however, came the problem of defining and using terms consistently and accurately. Thus the *CODATA Bulletin* Number 23 (1991) production of the "CODATA Glossary", with terms relating to data, data capture, and data manipulation, proved to be a valuable resource for facilitating the exchange of data in the international community.

Summary

Probably it was the computer that most impacted the focus and growth of CODATA during this period. Not only were computers becoming increasingly pervasive in the production and dissemination of data, they were being employed in exploiting and mining data to gain insights that otherwise would not have been discernible. Integration across disciplines became much more feasible. The expansion of the Internet made it possible to exchange vast quantities of data quickly and freely for analysis, comparison, and collaboration, with both positive and negative implications - positive in the sense that duplication of effort and costs were reduced; negative in the sense that the established means of validating the quality of those data were being by-passed. After many years of printing and disseminating its output products by conventional means, CODATA was sufficiently nimble to adapt its publication strategy regarding the *Bulletin* and Conference Proceedings.

With the approval of 16 new Task Groups during this third complete decade of its existence, CODATA's scientific activities were strong. National and Union Memberships were stable and finances were adequate to provide seed funding for worthy projects. Thanks to matching grants and thousands of hours of volunteer time, the value and impact of the scientific program far exceeded the amounts shown on the financial reports.

The 2001-2010 Decade

Governance

Montreal, Canada, was the site of the 23rd General Assembly in October 2002 with 22 National and Union Delegates participating. Having a lighter agenda than usual, the General Assembly had time to grapple successfully with a controversial Constitutional amendment addressing the issue of some National Members not paying their dues in a timely or complete manner. On another front, given the ‘greying’ demographics of CODATA, a very positive outcome was



Shuichi Iwata

Shuichi Iwata (Japan) as President, and Alexei Gvishiani (Russia) and Sun Honglie (China) as Vice Presidents. Paul Mezey was re-elected as Secretary General.



Alexi Gvishiani

A very active professor at the University of Tokyo with diverse interests ranging from nuclear engineering to design science to technology information, Iwata brought exceptional energy and breadth of data-related experience to the office. Gvishiani contributed a strong background in mathematics and earth science as attested by his many, varied international connections and responsibilities. Sun also had extensive experience in the earth sciences, his acumen having been recognized by a number of senior appointments in the Chinese Academy of Sciences and in ICSU.

As part of a strategy to revive national involvement in CODATA, the 24th General Assembly was held in Berlin,



Sun Hongli

Germany, attended by 23 National and Union Delegates. For the first time in CODATA's history, a Vice President chaired the deliberations. Due to illness, President Iwata was unable to make the journey so Vice President Gvishiani presided in his stead. One of the key decisions was to instruct the Executive Committee to initiate plans to draft a strategic plan that would address the issues raised in a recent ICSU Priority Area Assessment on Scientific Data and Information. Following up on the previous action regarding the involvement of young scientists, the General Assembly formally recommended that each National Member consider the establishment of Young Scientist Awards in their respective jurisdictions that would encourage the participation of scientists younger than 35 years at future CODATA conferences. Elections resulted in two new Officers: Krishan Lal (India) as a Vice President and Robert Chen (US) as Secretary General. Jean-Jacques Royer was re-elected as Treasurer.



Krishan Lal

Lal, a distinguished solid state physicist and former Director of India's National Physical Laboratory, was actively involved in promoting activities related to scientific data and their application in societal development, particularly in less-developed economies. He was acclaimed as one of the few active scientific leaders in India who were deeply involved with issues such as open access, the information society, and the ideals of open access to them.



Robert Chen

Chen, a geographer then serving as deputy director of Columbia University's Center for International Earth Science Information Network (CIESIN), had contributed to intergovernmental activities on climate change and served on panels and boards focussing on research data and information, as well as the CODATA Executive Committee. He thus brought an impressive amount of relevant international experience and connections to CODATA's leadership.

Returning to China after 14 years, the 25th General Assembly convened in Beijing in October 2006 with 29

Delegates representing National and Scientific Union Members. Much of the discussion centered around the draft 2006-2012 Strategic Plan and its recommendations, not the least of which was the approval of a new mission statement:

"The mission of CODATA is to strengthen international science for the benefit of society by promoting improved scientific and technical data management and use."

Another key decision was to establish and develop the *Scientific Data Across the Digital Divide Program* as a cross-cutting CODATA initiative focussed on improving the capacity of developing countries to access, collect, manage, and utilize scientific data in support of sustainable development. Reflecting ongoing concerns related to CODATA's financial stability, the General Assembly recommended that CODATA's dues structure be re-evaluated with a view to reducing barriers to membership on the part of developing countries, correcting imbalances in the current dues structure, and dealing with problems related to fluctuating currency exchange rates.



Steve Rossouw

The election cycle required the filling of three Officer positions. The new President was Krishan Lal and the two new Vice Presidents were Steve Rossouw (South Africa) and Gordon Wood. Lal and Wood had previously served as Vice President and Secretary General respectively, and Rossouw as an Ordinary Member of the Executive Committee. Rossouw, then at the University of South Africa, was an informatics/information specialist with a long history in CODATA including the Chairmanship of the South African CODATA National Committee along with service on the CODATA Executive Committee and numerous CODATA international committees.

Ukraine, a relatively new National Member of CODATA, having joined in 2000, hosted the 26th General Assembly in Kyiv in 2008, the first meeting in eastern Europe for

many years. Some 24 National Member and Union Delegates attended. One of the highlights of the General Assembly was the welcoming back of Australia and the United Kingdom as National Members following an absence of several years. A second highlight was the endorsement of the public review version of the 2006-2012 Strategic Plan slated for presentation at the forthcoming ICSU General Assembly, followed by final approval and dissemination by the CODATA Executive Committee. The General Assembly, having observed that the Young Scientist's activities had been one of the outstanding successes of the 21st CODATA International Conference held a few days earlier, strongly encouraged that the momentum generated by these activities be continued and built upon leading to the CODATA Conference in 2010.



Michel Sabourin

With respect to the election of Officers, Michel Sabourin (Canada) was elected Treasurer, the only vacant position that particular biennium. Having been the Delegate of the International Union of Psychological Science to CODATA for several years and having served for many years as the Treasurer of that Union, Sabourin brought not only experience in financial matters but was the first Officer from one of the social sciences, demonstrating how much CODATA had expanded its scientific span since its beginning.

Stellenbosch, in the wine country near Cape Town, South Africa, was the actual site of the 27th General Assembly in October 2010 with 26 National and Union Delegates present. Although a variety of CODATA-related meetings had previously taken place in Africa, this was the first time a General Assembly convened on the continent. It is interesting to note that one of the approved amendments to the Constitution allowed the CODATA General Assembly henceforth to approve new Co-opted Members rather than having to receive the approval of the ICSU Executive Board – a sign that CODATA was truly becoming master of its own house.

It is also of interest to observe the ongoing financial concerns, as another dues committee, just four years after the previous one, was to be appointed to analyze the dues structure and propose a structure that was “fair, objective, transparent and takes into consideration the impact of the current economic situations of the paying Members of CODATA”. From the perspective of 2010 it is clear that, somewhat like the United Nations, CODATA was still highly dependent upon the willingness and ability of Member Nations to pay their annual dues, being themselves subject to vagaries in their economies, their currencies, and the commitment of their decision makers to support international science.



Guo Huadong

With respect to new Officers, Guo Huadong (China) was elected President and Takashi Gojobori (Japan) and Fedor Kuznetsov as Vice Presidents. Secretary General Chen and Treasurer Sabourin were in mid-term. An earth scientist and past member



Takashi Gojobori

of the CODATA Executive Committee, Guo was Director-General of the Center for Earth Observation and Digital Earth in the Chinese Academy of Sciences. A recognized international expert in radar remote sensing, Guo had been one of the principal investigators in eight major international programs. Gojobori, a geneticist with a distinguished career in bioinformatics and a past member of the CODATA Executive Committee, was Vice Director of the Japanese National Institute of Genetics and a professor at the Center for Information Biology and DNA Data Bank of Japan.

Secretariat

During this decade, Executive Director Kathleen Cass administered the practical aspects of CODATA with distinction, innovation, and energy. Identifying and implementing a number of strategic ‘top-down’ activities, Kathleen was able to increase CODATA’s visibility and played an important role in CODATA successfully applying for major grants for new projects that otherwise could not have been undertaken using prevailing resources.

An auspicious development was that, after being located in the Hôtel de Noailles since 1974, the Secretariat moved in 2007 to a new location in Paris, near the Arc de Triomphe, along with ICSU and SCOPE (Scientific Committee on the Problems with the Environment). This historic building had been the location of countless CODATA meetings and many social events so that it was with some nostalgia that, following the Executive Committee meeting in March 2007, President Krishan Lal hosted a cocktail party to mark the occasion and bid adieu. However, in one sense, the move served both to celebrate the past and to herald a new chapter in CODATA’s story.

CODATA Conferences

The 18th Conference, *Frontiers of Scientific and Technical Data*, was held from 30 September to 3 October, 2002, in Montreal with about 250 participants from 37 countries. In a departure from past conferences, this one was jointly hosted by the Canadian and United States National Committees for CODATA with Gordon Wood and Harlan Onsrud of the University of Maine serving as co-Chairs. The program featured six plenary keynote lectures, along with 24 invited speakers, addressing the six cross-cutting conference themes:

- Preserving and Archiving of Scientific and Technical Data;
- Legal Issues in Using and Sharing Scientific and Technical Data;
- Interoperability and Data Integration;
- Information Economics for Scientific and Technical Data;
- Emerging Tools and Techniques for Data Handling; and
- Ethics in the Creation and Use of Scientific and Technical Data.

An innovation introduced at this conference was a special evening session open to the general public featuring two extraordinary lectures. The first was a bilingual presentation by Guy Baillargeon (Agriculture and Agri-Food Canada) on biodiversity and the *Global Biodiversity Information Facility*, the front-end to a collection of interoperable taxonomy and specimen databases. The second, by Fritz Hasler (NASA Goddard Space Flight Center) on *Visualizations of our Planet's Atmosphere, Land & Oceans* allowed the audience to experience a virtual reality fly-through, witnessing dust storms in Africa and smoke plumes from fires in Mexico – an illustration of how high-definition television was revolutionizing the way science could be communicated.



Arthur Pelton

Another conference highlight was the presentation of the 2002 CODATA Prize to Arthur D. Pelton of the École Polytechnique in Montreal, in acknowledgement of his innovative leadership in the field of computational thermodynamics as applied to metallurgical and chemical systems of interest in industrial and academic research. He and his group had led the development of the widely used and acclaimed *Facility for the Analysis of Chemical Thermodynamics (F*A*C*T)*, comprising evaluated databases of the thermodynamic properties for thousands of substances and solutions together with sophisticated software that used those data to calculate chemical equilibria information. At the time, *F*A*C*T* provided the most versatile tool currently available for thermodynamic analysis of industrial processes and materials design.

Closing the conference on a futuristic note was a session called CODATA 2015 during which three speakers in the fields of informatics, physics, and biology cast their vision of what their respective fields would look like in 13 years.

The Information Society: New Horizons for Science was the title of the 19th Conference held in Berlin, November 2004 with 260 participants representing 28 countries and several key international organizations in attendance. In response to current concerns in the data world, four plenary sessions addressed different important aspects of data – data and society, data archiving, data

visualization, and scientific data mark-up languages. Two of the distinguished keynote speakers were Johann Gasteiger of the University of Erlangen-Nuremberg in Germany and Yoshiyuki Sakaki of the Genomic Science Center of the Riken Institute in Japan. Gasteiger, a leading innovator in using databases for discovery in chemistry and drug design, spoke on *Design and Data in Chemistry: the Interplay of Ideas and Facts*. Sakaki, a world leader in using large scale genomic and proteomic databases to understand the strategy of life acquired through biological evolution, dealt with *Comparative Genomics: a Bioinformatics-based Key Approach for Understanding Our Genome*.

Reaching out to the broader scientific community, the conference also featured a stimulating session entitled *World Summit on the Information Society (WSIS): the Scientific Road to Tunis* which had taken place in Geneva in December 2003. Adama Sammeskeou, President of the WSIS Preparatory Committee for the Geneva Phase, was the keynote speaker, focussing on open access for all. Representatives from ICSU and UNESCO rounded out the session with presentations regarding their WSIS-related activities.



Jean Bonnin

In what was now becoming a feature of CODATA conferences, the CODATA Prize was awarded to Jean Bonnin, President of the *Earth Data Network for Education and Scientific Exchange (EDNES)*, in recognition of his outstanding achievements in the areas of geophysics, seismology, and geodynamics. Working in close collaboration with other French and American geologists and geophysicists, he had been an active member of the group that had created a revolutionary new Plate Tectonics model of the earth's structure in the 1960s. As the President of EDNES, he had carried out a variety of projects that integrated data activities in a number of European countries, projects that were also important to organizations in other countries, including Russia and the US. His distinguished career encompassed several fields of interest, including global tectonics and lithosphere deformation, data management, and seismicity.

Featuring the largest conference attendance in CODATA's history, with more than 600 participants from some 30 countries, the 20th CODATA Conference, *Scientific Data and Knowledge within the Information Society*, met in Beijing during October 2006. Presentations reflected the wide scope of data-centric activities in science and technology in which members of the CODATA family were involved. Of particular note were the keynote addresses by Xu Guanhua, China's Minister of Science and Technology, Jane Lubchenco, past-President of ICSU, Tony Hey, Vice President for Technical Computing, Microsoft, and Liu Depei, President of the Chinese Academy of Medical Sciences. The conference continued the emphasis on the Information Society initiated at the 2004 conference in Berlin as well as CODATA's involvement in the *World Summit on the Information Society* and related activities during the last biennium.

Running through the program was a sense of CODATA's self-evaluation of its purpose, demonstrated in a session on a retrospective view of its past 40 years together with a session looking at its future directions. A sample of the session topics illustrates the concerns of the time: Disaster Data, Computational Informatics, Role of Data in Promoting Public Understanding of Science, Solar-Terrestrial Data, e-Science, Virtual Observatories in the Geo-Sciences, International Polar Year activities, Chemical and Physical Data, Bioinformatics/Biodiversity, and Social Science Data Issues.

The 2006 CODATA Prize was awarded to John Rumble, a Past President of CODATA and, at the time, Technical Director for Information International Associates. A former Director of the Standard Reference Data Program at the US National Institute of Standards and Technology (NIST), he was honoured for being "an innovator in the world of scientific and technical data." He had been instrumental in extending data evaluation techniques in new disciplines, including engineering, materials, and biotechnology as well as for having been a long-time leader in using advanced information technology for developing computerized databases, online data networks, and data exchange standards. In every position held, he had promoted the development of quality standards and methods and the recognition of the importance of good data management and had worked

tirelessly to raise awareness of the importance of data to science and technology at both national and international levels. Very significantly, it was also noted that he had been instrumental in launching *CODATA's Data Science Journal*, the leading scientific publication in the field, and was the Journal's current Editor-in-Chief.

Kyiv, Ukraine, was the site of the 21st CODATA Conference, convened October 2008 under the theme *Scientific Information for Society – from Today to the Future*. The historic surroundings of the National Technical University of Ukraine easily accommodated the 400 participants. Consonant with its title, the conference emphasized the importance of engaging the younger generation of scientists who will be leading future developments in a world becoming increasingly dependent upon digital information and scientific data. Some of the sessions devoted to young scientists were among the best attended in the conference.

Following customary practice, the plenary sessions were designed to underscore the theme. Mikhail Zgurovsky, primary local organizer of the conference, gave the initial plenary presentation followed by an engaging talk by Bohdan Hawrylyshyn, at the time a board member of the World Academy of Arts and Sciences. Watura Iwamoto, representing UNESCO, described many of its initiatives in promoting the information society. Robert Jones from CERN, spoke of the European Union-financed EGEE project (Enabling Grids for e-Science in Europe) whose goal was to support the management and analysis of large amounts of data. Carole Goble, co-leader of the UK Information Management Group, outlined issues related to hardening and supporting the outcomes of the UK's e-science program.

Numerous parallel sessions reflected the multidisciplinary nature of the conference; a few examples must suffice to illustrate the range of topics covered: Information Society, Global Climate Change, Grid infrastructure, Biodiversity, Nanotechnology, Natural Disasters and Risk, International Polar Year, Biological and Genetics Data.

The 2008 CODATA Prize was presented to Liu Chuang from the Institute of Geography and Natural Resources in Beijing. She received the award in recognition of her



Liu Chuang

outstanding achievements on the international scene in scientific data research, along with policy and management issues, while simultaneously nurturing international cooperation and highlighting the many challenges facing data scientists in developing countries. More specifically she was recognized for working to improve data access and capacity building in the developing world both as co-Chair of the CODATA Task Group on Preservation and Access to Scientific and Technological Data in Developing Countries and as leader of the Global Alliance for Enhancing Access to and Application of Scientific Data in Developing Countries.

Nominally identified as having been held in Cape Town, South Africa, the 22nd CODATA Conference was held in October 2010 at the invitation of the South Africa National Research Foundation. As mentioned previously regarding the General Assembly, this was also the first time a CODATA International Conference had been held in Africa. More than 300 participants from all over the world made the journey to engage in this successful event as well as to enjoy the unique tourist opportunities.

Nadeli Pandor, South African Minister for Science and Technology, opened the conference with a particularly powerful speech addressing four key issues relevant to developing economies:

- access to science and technology by the poor and deprived communities;
- importance of intellectual property rights, taking into account the traditional and indigenous forms of knowledge;
- the need for international interdisciplinary networks to improve research and its applications and to address relevant socio-economic issues; and
- investment of government and scientific bodies in cutting-edge communication and storage technology.

Typical of a CODATA Conference, a wide variety of keynote addresses and parallel sessions were given. Of particular interest, given the venue, were those on e-Science and e-Infrastructure in many countries and regions; sustainable development; biological, environmental, and health data projects focusing on projects in Africa; and interoperability and access models.

To enhance the profile of the poster displays, two “One Minute Poster Madness” sessions were introduced – a stimulating new venture for a CODATA conference. After having their posters on display for the previous day, each author was given an opportunity for a one-minute précis of their poster with the three best presenters, based on five criteria, being awarded a prize.

Another notable feature of this conference was the enhanced emphasis on young scientists, 50 of whom from the African continent attended to discuss topics of particular relevance to them. Their discussions led to the *Cape Town Declaration* which was presented to the CODATA General Assembly, and each young scientist received a certificate commemorating their participation at the Conference. In another CODATA first, an award



was presented for the best Young Scientist Paper.

Young Scientists – 2010 Conference, Cape Town



Paul Uhlir

Paul Uhlir was awarded the 2010 CODATA Prize in recognition of his numerous contributions as a worldwide scholar on issues at the interface of science, technology, and law. Currently Director of the Board on Research Data and Information at the US National Academies, his work primarily focused on digital data and information policy and management, both nationally and internationally. Of particular relevance to CODATA's interests was his directing of the US Inter-Academy Panel's Program on Digital Knowledge Resources and Infrastructure in Developing Countries and his extensive

writing on digital commons issues and organizing of numerous meetings on the topic in many countries.

Scientific and Technical Activities

New Task Groups first approved in 2002-2010

Data Management and Virtual Laboratories (Later known as the **Virtual Laboratories in Earth Physics and Environmental Sciences** Task Group) – Recently developed measurement technologies had greatly increased the volume of global data in the areas of geothermodynamics, seismology, biology, satellite imagery, and linguistics. That increase in volume, combined with the interdisciplinary nature of many studies, required new data handling and knowledge extraction techniques. It was the goal of this Task Group, approved in 2002 after having been proposed by co-Chairs Herbert Kroehl and Jean Bonnin, to address those issues. Kroehl (US National Oceanic and Atmospheric Administration, Boulder, Colorado) was past Secretary General of the International Association of Geomagnetism and Aeronomy (part of the IUGG). Bonnin (winner of the 2004 CODATA Prize) was from the University of Strasbourg in France. One of their achievements was organizing a workshop on *Analytical Representation of Potential Field Anomalies* in Luxembourg.

Preservation & Archiving of S&T Data in Developing Countries (renamed to **Preservation of & Access to S&T Data in Developing Countries**) – Consistent with CODATA's increasing focus on data issues pertaining to developing countries, the primary aim of the Task Group, approved in 2002 and renamed in 2005, was to promote a deeper understanding of the needs of developing countries with regard to long-term preservation, archiving, and access to scientific and technological (S&T) data. Co-Chairs William Anderson of the US National Committee for CODATA and Liu Chuang, winner of the 2008 CODATA Prize, proposed a work plan that would advance the development and adoption of improved archiving procedures, technologies, open standards, and policies. The Task Group was to provide an interdisciplinary forum and mechanisms for exchanging information about S&T data archiving requirements and activities, focussing particularly on the needs of developing countries. A series of international

workshops and training workshops were developed in China (2003, 2005, 2007), South Africa (2002, 2005), Brazil (2007), Cuba (2008), Mongolia (2009) and Thailand (2009) together with the International Academy Panel and the UN Global Alliance for Information and Communication Technologies and Development. These activities raised the CODATA profile and played a very positive role in bridging the digital divide.

Biological Collection Data Access (renamed to **Observation & Specimen Record**) – The objective of this Group, initiated in 2002 and renamed in 2008, was to improve accessibility to existing and emerging biological collection databases at the international level by developing and coordinating proposals for data and metadata standards and standard protocols. At the time this knowledge base was largely under-utilized, because its highly distributed, heterogeneous, and complex scientific nature obstructed efficient information retrieval. Databases and networking were seen as the key to exploiting the potential value of biological collections for science, government, education, the public, and commerce. Walter G. Berendsohn (Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität, Berlin) and Stanley Blum (Center for Applied Biodiversity Informatics of the California Academy of Sciences) were the co-Chairs.

The modeling language developed by this Group was recognized as a significant contribution to the bioinformatics community by providing, for the first time, an accepted community standard for defining and accessing biological collections data. The Group played an important role in the development of a robust open access shared data infrastructure which is a critical requirement for the assessment of the impact of global changes in the environment. As of 2009, they had focussed on two projects: specific recommendations to the *Global Biodiversity Information Facility* regarding their observational data model; and a White Paper that described the opportunities for further ecological understanding and biodiversity conservation through the special characteristics of an expanded observational data model on species occurrence.

Anthropometric Data & Engineering – Representing another broadening of CODATA's interests by 2004, this Task Group had as its objective the creation of a system called *World Engineering Anthropometry Resource (WEAR)*. Anthropometry is the science of human body dimensions. The proposed system would include not only the latest three-dimensional surface anthropometric data from all over the world but also traditional anthropometric data, fit and accommodation information, analytic and software tools, and guidance or intelligent agents for using the information effectively. Co-Chairs Hein Daanen (TNO Human Factors, The Netherlands) and Sandy Ressler (National Institute of Standards and Technology, US) pointed out the potential benefits to many organizations worldwide, including, for example: hospitals, health statistics departments, and militaries, along with manufacturers of clothing, furniture, automobiles and airplanes. Along with holding two international workshops – Johannesburg (2005) and Rio de Janeiro (2005) – they developed a new structured naming method enabling standardized anthropometric data sharing. The beta version was implemented and demonstrated at the triennial *International Ergonomics Association* conference in Beijing, August 2010.

Low Dimensional Materials and Technologies Network – While considerable work had been done worldwide on databases addressing the macroscopic properties of materials, a need had been identified for databases dealing with the properties of materials in the range of 1-1000 nanometres where quantum effects are realized. Following their successful 2004 application, this Task Group planned to address that need by (a) identifying groups which could contribute to adding relevant databases on size effects in materials properties; (b) stimulating collection and organization of information on methods of modeling of size dependence of materials properties; (c) organizing connections between groups collecting data on size dependence of materials properties, and (d) creating a network of national and international organizations involved in work on these areas. Its Chair was Hwang Huey Liang, Tsing Hua University, a well-respected expert in this field. For reasons not made clear, the Group did not apply for renewal in 2006.

Data Sources for Sustainable Development in SADC (South African Development Community) Countries

– Proposed in response to one of the recommendations coming out of a 2005 CODATA workshop in Pretoria, this Task Group was formed in 2006 to determine the current status regarding the nature, volume, preservation, and accessibility of scientific and technical data for sustainable development in the SADC countries. This was to be accomplished by holding workshops and establishing liaison with relevant CODATA Task and Working Groups. Lulama Makhubela, with an extensive background in information science and service as Executive Director of Research at the National Development Agency, South Africa, and Conrad Sebego were the original co-Chairs.

Thirty-one delegates attended their initial workshop in Pretoria in May 2007, convened in collaboration with the ICSU Regional Office for Africa, the National Research Foundation of South Africa, the CODATA Task Group on Preservation of and Access to Scientific and Technical Data in Developing Countries, the South African National Committee for CODATA, and the United States National Committee for CODATA. A work plan of six concrete steps was developed.



Task Group on Data Sources for Sustainable Development in South African Development Communities

Comprehensive Information System on Natural Disaster Mitigation

– Quite different from the goals of Task Groups in earlier decades but reflecting the increasingly widening sphere of CODATA's interests, the objective of this Task Group was to work on the mitigation of major natural hazards and disasters by establishing a natural disaster database including historic and real-time components combined with a scientific/technical model

for disaster preparedness and mitigation in one or two developing countries or regions. Those reviewing the 2006 proposal agreed that data were a critical link for assessing disaster risks and that improving access to them was relevant, timely and important for mitigation and insurance. Co-Chairs were Wang Angsheng, Director of the Center of Disaster Reduction, China, and marine biologist Lars Hernroth of the Royal Swedish Academy of Sciences. The Group however did not continue past 2008.

Exchangeable Materials Data Representation to Support Scientific Research & Education –

Essentially, the goal of this Task Group endorsed in 2006 was to help the materials community catch up with other scientific communities that were advancing their endeavours through modern informatics, such as the use of distributed hardware systems, software, databases, and expert systems, as well as by exploiting the exponential growth in computing power, communication bandwidth, and storage capacity. Whereas the previous Materials Database Standards/Management Task Group, approved in 1988, focussed on the macroscopic (engineering) properties of existing materials, this Task Group planned to focus on the microscopic properties of substances relevant to the design of new materials for specific applications. To that end, they planned to promote open standardization and mapping of metadata along with general formats for data exchange and retrieval. Co-Chairs were Toshihiro Ashino (TOYO University, Japan) and Laura Bartolo (Kent State University, US). Members of the Group came from major materials database-providing organizations in the world and planned to promote collaboration among these organizations in order to establish an international information registry for materials science.

They met in 2008 in New Delhi where they focussed on: identifying materials information and knowledge resources; data format and semantic issues for sharing material information and knowledge; scientific commons; and integrating materials data into university science, technology, engineering, and mathematics education.

International Polar Year Data Policy & Management

Subcommittee – The objective of this Group was to define and guide the data policy, strategy, and overall management approach for the ICSU/WMO (World Meteorological Organization)-sponsored International Polar Year (IPY - 2007/8), facilitating international cooperation among the various IPY scientific projects. More specifically, plans in their 2006 proposal were to: (a) outline a strategy for IPY data management with specific goals based on the scientific questions formulated in the IPY Science Plan and which would meet the special needs of the social sciences, the Education, Outreach and Communications Plan, and the general public; (b) determine the organizational data flow structure (including consideration of procedures, time requirements, transmission and archival of data) required to support the science objectives; (c) assess the extent to which existing and proposed data management systems could be used by the IPY science projects; (d) provide expert advice to the IPY Joint Committee and also to endorse IPY projects on all IPY metadata and data management issues; and (e) to establish and provide ongoing guidance to the IPY Data Information Service along with a regular review of its effectiveness. Mark Parsons (National Snow and Ice Data Center/World Data Center for Glaciology, University of Colorado) and Taco de Bruin (Royal Netherlands Institute for Sea Research, The Netherlands) were the initial co-Chairs.

By 2010, a prototype Polar Information Commons had been developed to provide a forum for data providers to expose their data to the world and share them, without restriction, through open protocols on the Internet. Initial efforts included numerous presentations at meetings and workshops along with a web page describing the concept.

Global Information Commons for Science Initiative-EU (GICSI-EU) Activities –

An international initiative of CODATA and a multi-stakeholder project arising from the second phase of WSIS in Tunis in November 2005, this Task Group had the goal of accelerating the development and scaling up of open scientific data and information resources on a global basis, with particular focus on ‘common use’ licensing approaches. The aim of the Group, approved in 2008, was to build on previous exploratory activities and to develop an organizational and

research capacity in Europe in order to promote and consolidate the GICSI-EU Initiative. The initial themes to be addressed by the Task Group were: (a) impact of the current Intellectual Property Rights regime and public domain policies on the scientific communities; (b) alternative legal and technical regimes built by the scientific communities; (c) public research institutions' policies and practices for deposit of data and archiving of preprints and publicly funded research reports, dynamic distribution of metadata, and technical provision for preservation of the research record through standardisation of open formats and content migration; and (d) data sharing in EU-funded research, for example: database federation in EU Research Infrastructures, common ontologies for data mining and harmonization of privacy regulation for open access to content in biomedical research.

The importance of this activity to CODATA was indicated by the appointment of Paul David, Senior Fellow at both the Oxford Internet Institute and the Stanford Institute for Economic Policy Research, as a co-opted Member of the CODATA Executive Committee with primary responsibility to take the lead role.

In November 2007, CODATA became more formally aligned with COMMUNIA, a 36-member European project aiming to build a network of organizations to become the single European reference point for high-level policy discussion and strategic action for all issues relating to the public domain in the digital environment. The Task Group co-organized two COMMUNIA conferences. The first, in Belgium (July 2008), assessed the economic and social impact of digital public domain throughout Europe. The second, in Italy (June 2009), was attended by over 110 participants from more than 25 countries. They addressed the problems and solutions being developed to provide timely and lower cost access to scientific information in the published literature and the use and re-use of biophysical materials, research infrastructure facilities, databases and computational resources, and algorithms.

eGY Earth & Space Science Data Interoperability –

Reflecting ongoing concerns related to making disparate datasets widely available and readily usable, this Task Group had two main objectives: to promote the establishment and usage of modern and open data access standards and to help develop efficient means to achieve inter-operability across the Earth and Space sciences. Receiving approval in 2008, they sought to promote free access to metadata exchange and integration worldwide. Specific results anticipated were: development of a multi-disciplinary intelligent geographical information system (GIS) for the globe; creation of an electronic/hardcopy world atlas of the earth's magnetic field covering the period from 1500 to 2010; and the launch of a new geomagnetic data center for the Commonwealth of Independent States (CIS). Chair Alexei Gvishiani (Geophysical Center of Russian Academy of Sciences), assisted by Vice Chairs Eric Kihn (National Geophysical Data Center of the US National Oceanic and Atmospheric Administration) and Jean Bonnin, headed the 20-member, multi-national Group.

Their initial activity was co-organizing the International Electronic Geophysical Year (eGY) Conference *Electronic Geophysical Year: State of the Art and Results*, held June 2009 in Pereslavl-Zalessky, Russia. About 120 participants from eight countries (Russia, France, the US, Iran, Germany, Slovakia, Austria, and Ukraine) attended. The program focussed on the development of specialized virtual sources of geophysical data (virtual observatories) in the global network; improvement of systems of data selection, storage, and scientific analysis of data; and artificial intelligence methods in geoinformatics and geosciences.

Data at Risk – This Task Group proposed to complement CODATA's ongoing concerns about data quality and data accessibility with a plan to address the issue of data at risk of being lost – those data that predated the digital era and were recorded on media, such as photographic emulsions, paper, or magnetic tapes, which easily become fragile or obsolescent and are often discarded because they cannot be easily accessed or used. When such 'old' data are arbitrarily dismissed as unreliable, the loss to science, especially in fields involving long-term time sensitivity, is serious. As part of

their proposal approved in 2010, the Group sought to seek out and document those repositories of data considered to be at risk before the observations which they record have been preserved digitally for scientific purposes. The intended products were a preliminary assessment ranking the datasets in terms of preservation or rescue needs followed by a searchable online inventory in the public domain. Elizabeth Griffin (Herzberg Institute of Astrophysics of the National Research Council of Canada), and John Faundeen (Earth Resources Observation and Science Center of the US Geological Survey) were appointed as Chair and Vice Chair respectively.

Data Citation Standards and Practices – Very much in line with CODATA's objectives of accessibility to data, this Group was approved in 2010 to address new challenges arising from the electronic publishing of scientific data, in particular the need for mechanisms for citing online references to data in ways that can be relied upon for many years. Data citation standards and good practices can form the basis for increased incentives, recognition, and rewards for scientific data activities that in many cases were currently lacking in all fields of research. The rapidly-expanding universe of online digital data holds the promise of allowing peer-examination and review of conclusions or analyses based on experimental or observational data, as well as the ability for subsequent users to make new and unforeseen uses and analyses of the same data – either in isolation or in combination with other datasets – but that capability is severely compromised if those data and their provenance cannot reliably be found and determined.

Some specific issues to be addressed included: technical (interoperability, citation formats, and metadata); differing needs of scientific disciplines (granularity and cost implications); sustainability (registration and persistent identifiers); intellectual property rights; and citation and attribution norms. Heading the widely representative Group members were co-Chairs Bonnie Carroll (President, Information International Associates, US), Jan Brase (German National Library of Science and Technology, Germany) and Mark Thorley (Natural Environment Research Council, UK).

Global Roads Data Development – The objectives of this Group also demonstrate how the means by which CODATA fulfills its mission have evolved. Compared, for example, to the painstakingly evaluated, relatively few, data representing the properties of key chemicals – a foundational enterprise for CODATA in its early days – the data on roads were to be derived in copious amounts by a variety of new data collection technologies, including semi-automated extraction algorithms from high resolution satellite imagery, Global Positioning System (GPS) devices, and 'crowd sourcing' methods that take advantage of community contributions to data development. Continuing its preliminary work as a Working Group, the Task Group was authorized in 2010 to pursue the development of a roads data set catalogue and to release the initial version of the *Global Roads Open Access Data Set* (gROADS), backed up by plans to develop the infrastructure necessary for regular updates and releases. As an indicator of the importance of this issue worldwide, in addition to being a CODATA Task Group, this initiative was an approved task of the UN Global Alliance on ICT for Development of Open Access to and Application of Scientific Data in Developing Countries and an element of the Group on Earth Observations (GEO) work plan, and was endorsed by the Global Spatial Data Infrastructure Association and the GIS Corps of the Urban and Regional Information Systems Association. Group co-Chairs were Olivier Cottray (Information Management & Mine Action Programs, France) and Alex de Sherbinin (Columbia University, US).

It should be noted as well, in recalling part of CODATA's original *raison d'être* was to be an umbrella under which disparate bodies might work together, that this Group's affiliation with CODATA allowed their project to be viewed as a truly international, interdisciplinary scientific collaboration and not simply the venture of one or two solo organizations.

Other Developments of Interest

Fundamental Constants and the SI Units

As this Task Group celebrated its 40th birthday in 2009, recognized internationally as the definitive body for periodically carrying out least squares adjustments of the fundamental constants and issuing recommended values of the constants and conversion factors for energy equivalents for worldwide use, it was learned that its values will be used in the proposed changes to the International System of Units (SI). The recommendations of this Task Group to the Consultative Committee for Units (CCU) have contributed to the CCU's recommendation to the International Conference for Weights and Measures (CIPM), and through the CIPM to the General Conference on Weights and Measures (CGPM), that is expected to change the nature of the SI Units in a significant way. In particular, it is anticipated that the adoption of these changes will make it possible to eliminate the international prototype of the kilogram as the last of the physical artifacts, such as the international prototype of the metre, which was first accepted as the standard for that unit in 1889.

In retrospect, it is important to note the symbiotic effect that the work of this Task Group has had on precision measurements worldwide. As the precision and accuracy increased with which the Fundamental Constants could be determined from experimental data, scientists making those measurements were stretched to keep abreast, leading in some cases to the adoption of totally new techniques – an example being the volt which since the 1970s has been realized in terms of physical constants rather than the electromotive force of a group of standard cells.

Putting this in context, it is not an exaggeration to observe that the impact of this CODATA Task Group is enormous and at the core of the redefinition of the world's fundamental system of measurements. This is probably the most visible CODATA output, with tens of thousands of users consulting the website per month, and CODATA is clearly established as the definitive source for the recommended values of the fundamental constants.

Meanwhile, following its plan of releasing updated versions of the constants at least every four years, the

Task Group released adjustments in 2002, 2006 and 2010 supplemented by the publication of prominent peer-reviewed scientific journal articles documenting their methods and adjustments.

Secretariat Initiatives

Assuming a more proactive, strategic role in positioning CODATA's scientific agenda, the Secretariat successfully identified and implemented the following 'top-down' activities during this period.

International Symposium on Open Access

Another strong indicator of CODATA becoming increasingly recognized as a major player on the international scene was its organizing, in collaboration with UNESCO and ICSU, of the *International Symposium on Open Access and the Public Domain in Digital Data and Information in Science*, at UNESCO Headquarters in Paris, March 2003. Some 140 leading experts and managers came from government and academic sectors and from both developed and developing countries. All were involved in the creation, dissemination, and use of data and information in public research. They sought to describe the role, value, and limits of public domain and open access to digital data and information in the context of international research. Legal, economic and technological pressures involved were reviewed as well as the ways to preserve and promote public domain and open access to science and technology data and information on a global basis, with particular attention to the needs of developing countries.

Review of the Global Information Biodiversity Facility (GBIF)

GBIF is an international non-profit organization which provides free and universal access to data regarding the world's biodiversity, a project that originated at the Organization for Economic Cooperation and Development (OECD). Under its Memorandum of Understanding, GBIF was established for an initial period of five years with a comprehensive review mandated at the end of three years. Given CODATA's involvement and pool of expertise in scientific and data issues, CODATA

was chosen in 2004 to co-lead that review with KPMG, a global network of professional firms providing audit, tax, and advisory services in over 144 countries.

Information Commons for e-Science

Maintaining the momentum generated in the previous Symposium, CODATA and several partner organizations organized the *Information Commons for e-Science* workshop at the same Paris location in September 2005. The goal was to seek institutional policies and guidelines that would contribute to the creation of the ‘information commons’ for global e-science – a shared, international set of principles for public access to scientific data and information. In such a commons, access would be unfettered as much as possible from countervailing interests that tend to limit openness and cooperation such as national security concerns, intellectual property rights, and the practice of publicly funded researchers having periods of exclusive use of their data prior to publication of their findings.

World Summit on the Information Society (WSIS)

Of great interest to CODATA in this decade was the WSIS which took place in two stages – the first in Geneva in December 2003 followed by the second in Tunis in November 2005. Organized by the International Telecommunications Union under the patronage of the UN Secretary General, the WSIS was planned “to bring together Heads of State, executive heads of UN agencies, industry leaders, non-governmental organizations, media representatives and civil society ... to discuss the broad range of questions concerning the Information Society and move towards a common vision and understanding of this societal transformation”. In collaboration with ICSU, CODATA was very active in ensuring that science was clearly recognized as a key component of the Information Society.



**WSIS – Tunis –
Shuichi Iwata
presentation**

With respect to the initial stage in Geneva, CODATA was a formal partner for two major events: a scientific conference at CERN immediately prior to the Summit and a half-day high-level round table called

Science, Information Society and the Millennium Goals during the second day of the Summit. These events enabled the messages that ICSU and CODATA had been communicating throughout the preparatory process to be reinforced. An editorial entitled “Science and the Information Society”, written by then ICSU President Jane Lubchenco and CODATA President Shuichi Iwata, appeared in the prominent scientific journal *Science* in September 2003.

In 2004/2005 CODATA was extremely active in preparations for the second stage, in addition to the special session at the 19th CODATA Conference described earlier. Those achievements included:

- an inventory of more than 500 scientific activities that were related to the *Agenda for Action and Declaration of Principles* that came from the Geneva phase, prepared in collaboration with the US National Academies and the US National Committee for CODATA;
- another prominent editorial in *Science*, entitled “Science and the Digital Divide”, written by CODATA President Iwata and Secretary General Chen in October 2005; and
- an international workshop held at UNESCO Headquarters in Paris, co-sponsored with five other agencies, entitled *Creating the Information Commons for e-Science: Toward Institutional Policies and Guidelines for Action*.

CODATA's preparatory activities for the second stage at Tunis culminated in CODATA President Iwata's presentation to the eighth plenary session of the Summit. A particular highlight was his announcement of the launch, at a special WSIS Satellite Event earlier that week, of the *Global Information Commons for Science Initiative*. This multi-stakeholder undertaking had three main goals:

- improved understanding and increased awareness of societal benefits realized via easier access to and use of scientific data and information;

- wider adoption of successful models and methods for providing open, sustainable availability of publicly funded scientific data and information; and
- encouragement and coordination of the efforts of the many stakeholders in the world's diverse scientific community, in particular those in developing countries.

It is probably too early to judge the results of the WSIS but it is clear that CODATA played a key role in helping to achieve then-ICSU Deputy Director Carthage Smith's vision of WSIS being "the beginning of the end of the knowledge divide in science".

Evaluation of European Union Database Directive

Building on the work commenced in the previous decade, CODATA continued to monitor developments regarding intellectual property policy. In April 2006 on behalf of ICSU, CODATA submitted comments to the European Commission on the EU Database Directive in response to the evaluation of the Directive that was published by the Commission in December 2005.

Group on Earth Observations (GEO) Summit

As further evidence of becoming an influential international body wherever scientific data issues are involved, CODATA played a very active role in the GEO Ministerial Summit held in Cape Town, South Africa in November 2007. Desiring to establish a Global Earth Observation System of Systems (GEOSS), the GEO had requested CODATA to prepare a White Paper on implementation guidelines for the data sharing principles of GEOSS. Drafted by an impressive international team of CODATA-related authors, the White Paper was well received at the GEO Plenary meeting associated with the Summit. Active cooperation continued at the *GEO-VII Plenary and Ministerial* in Beijing in November 2010 where the delegates approved the Data Sharing Action Plan prepared by the GEOSS Data Sharing Task Force, which included CODATA Prize winner Paul Uhlir and Secretary General Robert Chen as co-chairs representing

ICSU. CODATA organized well-attended side events at the GEO Plenaries in Cape Town (2007), Bucharest (2008), Washington DC (2009), and Beijing (2010) and also showcased its GEO-related work at CODATA and other prominent scientific conferences.

Polar Information Commons

CODATA, working with its International Polar Year Data Policy & Management Subcommittee and other international organizations, launched the Polar Information Commons (PIC). During January-June 2009, CODATA had held a number of meetings and adopted an outreach strategy, noted by a reference in *Nature*, leading to the launch of the PIC in Oslo, Norway, June 2010. The

PIC booth was visited by many dignitaries throughout the week, including HRH Crown Prince Haakon of Norway and HSH Prince Albert II of Monaco.



Polar Information Commons Booth, Norway 2010 - David Carlson (Executive Director IPY Program Office) and HSH Prince Albert II of Monaco

Information Technology Activities

Compared to previous decades, work on new data and information technologies played more of an incremental than a pivotal role. The increased speed and reach of the Internet, combined with the remarkable processing power and memory available at the desk of virtually every scientist worldwide, removed many of the latent technical barriers to data access, bridging to a certain extent the 'digital divide'. As described above, the WSIS events perhaps summarize best the impact of information technologies on CODATA.

Publications and Other Outputs

Data Science Journal

CODATA embarked on another new journey with the launch of the electronic *Data Science Journal* in 2001. Part of the rationale was the realization that dissemination of CODATA products had been problematic over the years. Much good work in the form of Bulletins, Conference Proceedings, Reports, and monographs had been published, but not always in a widely accessible manner. In particular, the work had not been consistently noticed by international abstracting and indexing organisations that were now the primary gateway to the literature for virtually every researcher. *The Data Science Journal* was to be a fully-recognised periodical, a high-quality, peer-reviewed journal on data science reaching interested scientists throughout the world with every expectation that it would be abstracted and indexed by a number of secondary publishers. Furthermore, it was to be an open-access journal, an online-based publication with free downloading and dissemination, financed by CODATA and grants, most recently, from UNESCO.

Other key reasons cited for starting the Journal are instructive:

No existing data science journal – No international journal existed that covered the topics associated with the collection, evaluation, analysis, dissemination and use of scientific data, especially on a multi-disciplinary basis. The proposed Journal would be a unique publication facilitating the development and sharing of new knowledge about all aspects of scientific and technical (S&T) data.

Need to share Data Science results across disciplines – from a number of perspectives, S&T data activities have many common features such as their structure, relationship, display, and use. As new ideas are developed in data science and informatics it is important for the progress of science in general that researchers, regardless of specialities, are aware of the latest advances.

A journal gives identity to a discipline – Data Science was seen to be emerging as an important component of the information revolution. The new journal would provide an international focal point for this work and experience had

shown that, ultimately, a journal allows individual researchers to feel part of, and to nurture, a scientific community.

Service to Member countries – CODATA must bring value to its National Members and to date that had been done through Task Group work, workshops, conferences, and monographs. A periodical would be another means of providing value.

Overall visibility – In times of increased competition for recognition and resources, CODATA needed to consider maximizing its visibility. The journal would inform subscribers of the value and importance of CODATA.

Increased access to Data Science by Member countries – At least a quarter of CODATA Member countries find it difficult to access literature because of high subscription rates, developed country-level book prices, expensive mail, and weak libraries. The journal would provide an important mechanism for access to current work on S&T data.



Jack Smith – First Editor of the Data Science Journal

Jack Smith (Queens University, Belfast) was appointed initial Editor-in-Chief, a favourable combination of his skills with the existing capacity for electronic publishing at Queens. In January 2006, after an open call to CODATA members for a new host institution and editor, John Rumble was

appointed Editor-in-Chief and the *Journal* was relocated from the University of Belfast to the Japan Science and Technology Information Aggregator (J-Stage), an electronic journal portal, where it remains and continues to serve the data science community. In 2009, Shuichi Iwata replaced Rumble as the new Editor-in-Chief.

CODATA Newsletter Goes Online

After many years of being printed and shipped around the world, the *Newsletter* was converted to a pdf format in April 2003 that subscribers could download from the CODATA website. In addition to expediting the publication process, this innovation saved printing costs at the CODATA Secretariat and distribution costs to the various National Members which had been responsible for its distribution within their jurisdiction. Some copies continued to be printed, however, for publicity and related purposes.

Summary

During this period, CODATA matured into a more confident, agile organization, quick to embrace new technologies and extend its reach across the so-called digital divide, to Africa in particular. That CODATA played a significant role in major worldwide programs like the WSIS, the IPY, and GEO was tangible evidence of the

international stature that has been achieved. The ability to attract substantial grants, combined with being alert to cooperation and leveraging opportunities, allowed CODATA's scientific accomplishments to surpass impressively those which would be considered to be in proportion to the financial resources available from National Members.

The approval of 14 new Task Groups during this period demonstrated not only that CODATA's scientific activities were vigorous but also that a corporate willingness existed for matured projects to terminate and for new, often risky projects to be undertaken. National and Union Memberships continued to be stable and finances were adequate to cover operations and provide seed funding for new activities. As in past decades, untold thousands of hours of volunteer time on the part of highly skilled experts allowed the value and impact of the scientific program to far exceed the nominal budget.

CODATA in Retrospect

When CODATA was established 45 years ago, terms like “database” and “data management” had not yet entered the vocabulary. A few dedicated scientists, often at the expense of their research careers, were engaged in compiling, organizing, and evaluating data in their own discipline, but no organized mechanism existed for communication with those in other disciplines who were engaged in similar pursuits. The use of computers to manage large collections of data was in its infancy. CODATA’s evolution over this 45 year period has tracked closely with the revolution in science that has made the intelligent management of data a crucial factor in solving complex scientific and technical problems. The projects detailed in this history illustrate the many ways in which CODATA has responded to changing needs in the scientific community.

One of CODATA’s greatest accomplishments has been to support the cross-fertilization of technology advances among diverse disciplines. In today’s world, in which almost every scientist is required to be a data specialist, CODATA remains relevant as the only organization that is addressing data issues on a multidisciplinary basis. It has fostered international collaboration in carrying out data-related programs, even during periods of political confrontation. Furthermore, CODATA has shown special interest in spreading the methodology of data management to developing countries. In all its activities, CODATA has worked to leverage its limited financial and human resources by developing joint projects with a multitude of other scientific organizations. This has been achieved with a minimum of bureaucracy supported by a small proactive Secretariat.

The challenges of today are different, and in many ways more complex, from those of the 1960s, but CODATA is well-placed to make a significant contribution. It has already taken steps to address issues such as open access, intellectual property rights, and long-term preservation of data. CODATA can use its experience in data evaluation to address the major issue of data quality. The Internet has facilitated easy access to massive amounts of scientific and technical data of widely varying quality, but a data user is often left with the problem of deciding which value to choose out of the many that are retrieved by an Internet search. Political controversy over global climate change is only one example of the need for verified, authoritative data sets that are relevant to pressing world problems. This is an opportunity for CODATA, with its multidisciplinary and multinational resources, to make a real impact. We can expect CODATA to continue to play a major role in science and technology in its next 45 years.

Appendix A

Sites of CODATA General Assemblies and International Conferences

Year	Location	General Assembly	Conference
1966	Paris, France	First	
1967	Moscow, USSR	Second	
1968	Arnoldshain, Germany	Third	First
1969	Rome	Fourth	
1970	St. Andrews, UK	Fifth	Second
1971	Washington, DC, USA	Sixth	
1972	Le Creusot, France	Seventh	Third
1973	Stockholm	Eighth	
1974	Tsakhkadzor, Armenia	Ninth	Fourth
1976	Boulder, USA	Tenth	Fifth
1978	Santa Flavia, Italy	Eleventh	Sixth
1980	Kyoto, Japan	Twelfth	Seventh
1982	Jachranka, Poland	Thirteenth	Eighth
1984	Jerusalem, Israel	Fourteen	Ninth
1986	Ottawa, Canada	Fifteenth	Tenth
1988	Karlsruhe, Germany	Sixteenth	Eleventh
1990	Columbus (Ohio), USA	Seventeenth	Twelfth
1992	Beijing, China	Eighteenth	Thirteenth
1994	Chambéry, France	Nineteenth	Fourteenth
1996	Tsukuba, Japan	Twentieth	Fifteenth
1998	New Delhi, India	Twenty-first	Sixteenth
2000	Stressa, Italy	Twenty-second	Seventeenth
2002	Montreal, Canada	Twenty-third	Eighteenth
2004	Berlin, Germany	Twenty-fourth	Nineteenth
2006	Beijing, China	Twenty-fifth	Twentieth
2008	Kyiv, Ukraine	Twenty-sixth	Twenty-first
2010	Cape Town, South Africa	Twenty-seventh	Twenty-second

Appendix B

National Members – Years of Membership in CODATA

Decade	1960s		1970s					1980s					1990s					2000s					2010s
Nation	6	8	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0
Australia																							
Brazil																							
Canada																							
Cameroon																							
Chinese Academy of Sciences																							
Chinese Academy in Taipei																							
Czech Republic																							
France																							
Georgia																							
Germany (FRG)																							
Germany (GDR)																							
Hungary																							
India																							
Indonesia																							
Ireland																							
Israel			71																				
Italy			71																				
Japan																							

Decade	1960s		1970s					1980s					1990s					2000s					2010s
Nation	6	8	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0
Netherlands			71																				
Nigeria																							
Poland		69																					
Senegal																							
South Africa																							
South Korea																							
Sweden			71																				
Switzerland																							
Thailand																							
Ukraine																							
United Kingdom																							
United States																							
USSR/Russia																							

Note - * - Associate Member

Appendix C

Scientific Union & Co-opted Members - Years of Membership in CODATA

Decade		1970s					1980s					1990s					2000s					2010
Scientific Union	68	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0
Astronomical (IAU)																						
Biochemistry & Molecular Biology (IUBMB)																						
Biochemistry (IUB)																						
Biological Science (IUBS)	69																					
Crystallography (IUCr)	67																					
Geodesy & Geophysics (IUGG)																						
Geographical (IGU)	67																					
Geological Sciences (IUGS)	67																					
Immunological Societies (IUIS)																						
Microbiological Sciences (IUMS)																						
Nutritional Sciences (IUNS)		73																				
Pharmacology (IUPHAR)																						
Psychological Science (IUPsyS)																						

Pure & Applied Biophysics (IUPAB)																						
Pure & Applied Chemistry (IUPAC)	67																					
Pure & Applied Physics (IUPAP)	67																					
Theoretical & Applied Mechanics (IUTAM)	67																					

Co-Opted Members	68	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0
Federation of Astronomical & Geophysical Services (FAGS)																						
International Council of Scientific & Technical Information (ICSTI)																						
International Society of Soil Scientists																						
World Data Centers (WDC)																						
World Data System Scientific Committee (WDS SC)																						
World Federation for Culture Collections (WFCC)																						

Appendix D

Task Groups and Commissions – Years of Service to CODATA

Decade		1970s					1980s					1990s					2000s					2010
Name	68	0	2	4	6		0	2	4	6	8	0	2	4	6	8	0	2	4	6	8	0
General Data Issues																						
Accessibility and Dissemination of Data																						
Artificial Intelligence and Computer Graphics																						
Asian-Oceanic Data Sources																						
CODATA and the World Wide Web																						
CODATA Referral Data Base																						
Comparative Mathematical Methodologies of Data Handling and Knowledge Extraction																						
Computer Use																						
Data Access Commission/ ICSU/CODATA Group on Data and Information																						
Data at Risk																						
Data Citation Standards and Practices																						
Data Sources for Sustainable Development in SADC Countries																						

Appendix E

CODATA Officers

Year	President	Vice President	Vice President	Secretary General	Treasurer
1966	Rossini	Vodar	Klemm		
1967	Rossini	Vodar	Klemm		
1968	Rossini	Vodar	Sutherland	Klemm ³	
1969	Rossini	Vodar	Sutherland	Klemm ¹	
1970	Vodar	Styrikovich	Sutherland	Klemm ¹	
1971	Vodar	Styrikovich	Sutherland	Klemm ¹	
1972	Vodar	Styrikovich	Jones	Egle ¹	
1973	Vodar	Styrikovich	Jones	Westrum	Kurti
1974	Melchior	Plebanski	--	Westrum	Kurti
1976	Melchior	Plebanski	--	Westrum	Kurti
1978	Kotani	Plebanski	Sytchev	Westrum	Kurti
1980	Kotani	Dubois	Sytchev	Westrum	Watson
1982	Hutchison	Dubois	--	Lide	Watson
1984	Hutchison	Dubois	Bylicki	Lide	Watson
1986	Lide	Dubois	Bylicki	Bussard	Watson
1988	Lide	Gurvich	Bylicki	Bussard	Crease
1990	Abir	Gurvich	--	Wood	Crease
1992	Abir	Gurvich	Tsugita	Wood	Crease
1994	Dubois	Gurvich	Tsugita	Wood	Crease
1996	Dubois	--	Kuznetsov	Wood	Chinnery
1998	Rumble	Tsugita	Kuznetsov	Mezey	Chinnery
2000	Rumble	Tsugita	--	Mezey	Royer
2002	Iwata	Gvishiani	Sun	Mezey	Royer
2004	Iwata	Gvishiani	Sun	Chen	Royer
2006	Lal	Rossouw	Wood	Chen	Royer
2008	Lal	Rossouw	Wood	Chen	Sabourin
2010	Guo	Gojobori	Kuznetsov	Chen	Sabourin

³ Secretary-Treasurer

Appendix F

First International CODATA Conference

on

The Generation, Collection, Evaluation & Dissemination of Numerical Data for Science and Technology

Akademie Arnoldshain, Germany, June 30-July 5, 1968

Country	Participant	Affiliation
Belgium	Louis Deffet	Institut Belge des Hautes Pressions
	Jean D. Drowart	Université de Bruxelles
	P. Hestermans	Institut Belge des Hautes Pressions
	Jean-Pierre Poupko	Bibliothèque Royale de Belgique
Canada	R. Norman Jones	National Research Council of Canada
France	Ibrahim Ansara	CNRS
	Christian Barbe	L'Air Liquide – Laboratoires Centraux
	Jean-Claude Chenavas	CNRS
	A.T. David	Institut Français des Combustibles et de l'Energie
	Georges Denègre	Direction Recherches et Moyens d'Essais
	Pierre Desré	Laboratoire de Métallurgie et Electrometallurgie
	Pierre Duffaud	Cie de Saint-Gobain, Direction des Recherches
	M. Grandclaude	Centre de Recherches Petrographiques et Geochimiques
	Jean Lochard	AFNOR
	Boris Vodar	Laboratoire des Hautes Pressions
	Robert Abbel	Institut für Dokumentationswesen
	Dieter Behrens	DECHEMA
	Karl-Christian Buschbeck	Gmelin Institut f. Anorg. Chemie
Germany	Robert Fugmann	Farbwerke Hoechst AG
	Otto Gekeler	AEG-Telefunken
	Peter Haas	IBM Germany
	Karl-Heinz Hellwege	Technische Hochschule Darmstadt
	Heinrich Kaiser	Institut für Spektrochemie und angewandte Spektroskopie
	Willy Kattwinkel	IBM Germany
	Wilhelm Klemm	Anorg. Chem. Institut der Universität Münster.
	Ottmar Knacke	Lehrstuhl f. Metallurgie der Kernbrennstoff u. Theoret. Hüttenkunde d. T.H. Aachen
	Fritz Krückeberg	Institut f. Angew. Mathematik der Univ. Bonn
	Heinz Lechmann	Bundesministerium f. wissenschaftl. Forsch
	Walter Lippert	Gmelin-Institut f. Anorg. Chemie
	Walter Oberender	Battelle Institut e.V.
	Ullrich Schley	Physikalisch-Technische Bundesanstalt
	Josef-Johannes Schmidt	Kernforschungszentrum Karlsruhe

	Klaus Schneider	Zentralstelle f. maschin. Dokumentation
	Manfred Schönberg	Farbwerke Hoeschst AG
	Eva-Dora Schröder	Zentralstelle f. maschin. Dokumentation
	Friedhelm Schulte-Tigges	Deutsches Rechenzentrum
	Barbara Starck	Physikalisches Institut d. Univ. Freiburg
	Herbert Stussig	Arbeitsgemeinschaftl. Industrieller Forschungsvereinigungen
Israel	I. Eliezer	Tel-Aviv University
	Aviezer S. Kertes	The Hebrew University of Jerusalem
Japan	Masao Kotani	Osaka University
	Yo-ichiro Mashiko	The Japanese Government Chemical Industrial Research Institute of Tokyo
	Takehiko Shimanouchi	University of Tokyo
Netherlands	H. P. Hogeweg-de Haart	Royal Netherlands Academy of Sciences
	W. M. Smit	Royal Netherlands Academy of Sciences
Poland	Andrzej Bylicki	Institute of Physical Chemistry
Sweden	Stig Sunner	Lund University
U.K.	Selby Angus	Imperial College of Sci. & Technology
	Gordon Black	The National Computing Centre
	R. E. Fairbairn	Office of Scientific and Technical Information
	C. G. Giles	Ministry of Technology
	E. F. G. Herington	National Physical Laboratory
	K. G. Heyden	Heyden & Son Ltd.
	Olga Kennard	University Chemical Laboratory, Cambridge
	O. Kubaschewski	National Physical Laboratory
	D. M. Newitt	Imperial College
	Malcolm H. Rand	Chemistry Division, A.E.R.E.
	R. W. Selwyn	Materials Data Ltd.
	D. J. Shields	Imperial Chemical Industries Ltd.
	Henry A. Skinner	University of Manchester
	Sir Gordon Sutherland	Emmanuel College, Cambridge
U.S.A.	Freeman F. Bentley	Air Force Materials Laboratory
	Edward L. Brady	National Bureau of Standards
	John T. Edsall	Harvard University
	Robert Fristrom	The Johns Hopkins University
	Norwood B. Gove	Oak Ridge National Laboratory
	Alexander G. Hoshovsky	COSATI
	Ralph Hultgren	University of California
	David R. Lide, Jr.	National Bureau of Standards
	Warner M. Linfield	IIT Research Institute
	Ellis R. Lippincott	University of Maryland
	Joseph F. Masi	Air Force Office of Sci. Research (SREP)

	Hendrik van Olphen	National Academy of Sciences
	Frederick D. Rossini	University of Notre Dame
	Stephen A. Rossmassler	National Bureau of Standards
	Philip Sadtler	Sadtler Research Laboratories, Inc.
	Roger G. Simard	American Society for Testing Materials
	Charlotte M. Sitterly	National Bureau of Standards
	Daniel R. Stull	Dow Chemical Company
	Yeram S. Touloukian	Purdue University
	Donald D. Wagman	National Bureau of Standards
	Robert C. Weast	Chemical Rubber Co.
	M. Kent Wilson	National Science Foundation
	Bruno J. Zwolinski	Texas A & M University
U.S.S.R.	L. V. Gurvich	Institute of High Temperature
	Victor N. Kondratiev	Academy of Sciences of the USSR
	V. Medvedev	Institute of High Temperature
	E.E. Spylrain	Institute of High Temperature
	M.A. Styrikovich	Academy of Sciences of the USSR
International Organizations		
	Lef Hjärne	International Atomic Energy Agency
	Guy Waddington	CODATA
	Christoph Schäfer	CODATA

