

The Fuel for 21st Century Science and Technology

The importance of data in science and engineering today is difficult to exaggerate. Data today are rapidly becoming the source of new discovery for tomorrow. Every scientist and engineer must now be familiar with the tools to collect, store, evaluate, analyze, visualize, archive and disseminate data. These tools complement traditional theoretical and experimental tools that we were raised on. Let me explain more about what I mean by these statements, and how they are relevant to CODATA today.

March 2002

Highlights

- 1 (2)** The Fuel for 21st Century Science and Technology
- 2 (3,7)** Upcoming Meetings
- 4 (5)** General News
- 5 (6-8)** National Members Section
- 8** CODATA Calendar
Brain Teaser page 6 !

<http://www.codata.org>

The Committee on Data for Science and Technology (CODATA) was established in 1966 by the International Council for Science (ICSU). Working on an interdisciplinary basis, CODATA seeks to improve the quality, reliability, information and knowledge processing, management and accessibility of data of importance to science and technology.

Science and engineering of the last four centuries has focussed on the need to reduce complex phenomena to their basic elements. Whereas a liter of air has about 10^{23} atoms in it, when we studied its behavior, we mostly looked at the details of the interactions of two or three atoms at a time. This reductionism led to the development of many powerful experiment tools that separated out individual small systems from much larger collections. Perhaps the ultimate example is the large particle accelerators that capture individual collision tracks among millions of other collisions. Theoretically also we have developed a suite of theories and computational procedures that model individual interactions, or the structure of individual systems.

The excitement and discovery of these processes will continue into the indefinite future as backbones of sciences. However, a new element has been added. The aggregation of individual results coming from this research is being captured and exploited using modern data technology. In turn, the large collections of data are rapidly becoming sources of new discovery, the basis for new scientific insight and understanding. Just as the theoretical and experimental tools developed in one discipline are applied to another – electronic, lasers, chemical separation, etc. – the tools for data discovery are applicable to every area of science and engineering.

So what does this have to do with CODATA? Everything! CODATA can be an important source of learning about data technology. It can be the place where new ideas and techniques are described and shared, not just in one discipline, but across disciplines. Data technology is remarkably broad in its utility, and it is a technique that can be directly applied across disciplines.

- CODATA must be a place where you can learn what is new in data technology and how to use it
- CODATA must be a place where you can network with peers from your discipline and other disciplines working with data technology
- CODATA must be a place that helps foster new data technologies

We do much of this now in a variety of ways.

- International conferences
- Specialist workshops
- Conference and workshop proceedings
- The new CODATA Journal
- Task and Working Groups
- Regional meetings



continued on page 2



continued from page 1

While these are moderately successful, the audience for our efforts is much larger than we are reaching. To do better, we must ask ourselves some important questions.

- How do we reach people (especially in face-to-face settings) while minimizing international travel?
- How do we expose people in one discipline to new technology from another discipline in terms they understand?
- How do we provide instructional material on an introductory level, on an expert level?
- How do we take advantage of the ubiquitous nature of the Web to share knowledge and teach?

CODATA must make a commitment to become a leading source of knowledge about modern data technology. This is the soul of our organization. What will it take?

More effectively capturing and disseminating the knowledge and information we currently share through existing CODATA activities.

Identifying important areas of data technology and creating new instructional material aimed at introductory and expert users.

Partnering with other groups to develop new data technology – standards, dissemination, analysis – where there is a need – virtual laboratories are an example.

Intensifying our efforts to ensure that national and international policies do not become barriers to progress in data work.

The unifying theme through all these challenges is the pervasive nature of networking. We can create and share through the Web in ways that were not possible through the printed page. CODATA must make special efforts to take advantage of the Web, whether through routine sharing of information – newsletters and announcements, videoconferencing to facilitate meetings, Webcasts of lectures and talks, and tutorials that can be easily downloaded.

In October 2002, the CODATA General Assembly will be meeting to elect new officers and set up new Task Groups to meet these challenges. I strongly believe CODATA is ready to move forward in these areas, but our success will come only from a recognition that we must work together, be innovative and take chances.

John Rumble, Jr
President CODATA

UP COMING

18th International CODATA Conference

CODATA 2002, entitled *Frontiers of Scientific and Technical Data*, is to be held in Montreal, Canada from 29 September to 3 October 2002. The scientific program for CODATA's 18th Conference is being organized around six cross-cutting themes:

- Data Archiving
- Interoperability and Data Integration
- Emerging Tools and Techniques for Data Handling
- Legal Issues in the Use of Scientific and Technical Data
- Information Economics for Scientific and Technical Data
- Ethics in the Use of Scientific and Technical Data

Each theme will be introduced by a Keynote speaker in a plenary session and thereafter explored in a parallel session from the perspective of four different disciplines.

The Scientific Program Committee is pleased to report that plans for these thematic sessions, along with some 30 discipline-specific contributed sessions, are proceeding well and that significant progress has been made. Keynote speakers for four of the six thematic plenary sessions have agreed to participate and the remaining two speakers are in the final stages of being invited. In addition, 16 of the 24 speakers for the thematic parallel sessions have accepted their assignments and most of the others have been identified. Some 25 proposals to organize contributed sessions in a variety of specific areas are already in hand.

A novel addition will be a special plenary session, **CODATA 2015**, featuring three specialists sharing their vision of where data activities are going. Paul Ginsparg, a well-known innovator in information sharing, he is scheduled to be the Keynote Speaker.

Readers are invited to the CODATA web site <http://www.codata.org> for more details including the Call for Papers and forms for Conference and hotel registration.

G. H. Wood
Co-chair Scientific Program Committee



MEETINGS

The 1st CODATA Workshop on Archiving of Scientific and Technical (S&T) Data

A workshop on archiving of S&T Data will take place at the National Research Foundation in Pretoria, South Africa on 20-21 May 2002. It is been organized by the South African National Committee for CODATA, with the assistance of the CODATA Working Group on Data Archiving.

The workshop will be organized around four major themes, i.e. scientific issues, technical issues, management issues, and policy issues. International experts on archiving scientific and technical data will deliver keynotes on these themes followed by discussions by the participants. Academic, research, government and private organizations who have an interest and involvement in data archiving will be invited to send representatives to the workshop to share their experiences in the field with workshop participants.

The core group of experts will discuss the theme topics and will then open the discussion to a second level of attendees who will be afforded the opportunity to gain knowledge on what is becoming a "hot" topic in the field of archiving S&T data.

Although attendance and participation in the workshop will be mostly by invitation, a limited number of "volunteer" participants will be considered. All participants will be responsible for their own travel and accommodation expenses. There will be no registration fees.

More information on the workshop can be obtained from the co-chairs of the CODATA Working Group on Archiving as follows: Dr William Anderson band@acm.org or Prof Steve Rossouw steveros@iafrica.com. Details concerning the venue, accommodation and local transport can be obtained from Ms Henda van den Berg:

jhmvdberg@nrf.ac.za

Steve Rossouw
Co-chair of the CODATA Working Group
on Data Archiving

13th DSAO Task Group Workshop

The 13th Workshop on the Survey of Data Sources in Asian-Oceanic Countries (DSAO) CODATA Task Group will be held at the International Conference Room, of the Korean Institute of Science & Technology Information (KISTI), July 5th-6th in Daejeon, Korea.

The main theme of the workshop is "**Archiving Scientific Data in Asian-Oceanic Countries**". Presentations and discussions are expected on the broad field advanced theory and technology of information processing and data handling. Such issues concern the data sharing Asian-Oceanic countries, and other topics on factual data of various fields.

Some of the topics covered are

1. Methodologies and Application of Metadata
2. Data Visualization
3. Biological Science Data including, Bio-diversity, Proteomics and Bioinformatics
4. Geological, Data-Earthquake, Mine etc
5. Metrology Data
6. Other related Factual Data, Bio-hazardous Material, Infrastructure Protection.

The DSAO Task Group meeting will also take place in July in Korea and has been initiated by the Korean CODATA/DSAO member. All the expenses, airfare, accommodation, meal, domestic travel fare for all paper presenters and DSAO committee members are covered by the Korean Organizers who include Dr. Young Hwa Cho, President, (KISTI) and Prof. Mu-Shik Jhon, Korean Advanced Institute of Science and Technology (AASA). More information on the meeting is available on

<http://codata.kisti.re.kr/dsao2002>

Akira Tsugita
Chair DSAO



GENERAL NEWS

AROPA Workshop on Analytical Representation of Potential Field Anomalies for Europe

The AROPA workshop was organized by the European Center for Geodynamics and Seismicity (ECGS) in cooperation with the CODATA Task Group "Data Management and Virtual Laboratories". It was held in the "Institute d'Europe", Castle of Münsbach, Luxembourg, October 23-27, 2001.

The Scientific Organizing Committee was chaired by V.N. Strakhov (Russia), H.J. Götze (Germany) and B. Ducarme (Belgium). The secretary was U. Schäfer (Germany). The local organizing Committee was chaired by O. Francis (Luxembourg, ECGS secretary), and members included N. d'Oreye (Luxembourg), P. Codran (Luxembourg).

The objectives of the meeting were to demonstrate and discuss research results and, examine the potential of the new innovative interpretation techniques relevant to both the scientific and application communities and to promote the transfer of knowledge, technique and data concerning geopotential field representation. Besides the presentation of new theoretical and methodological elaborations, emphasis was also given to practical applications to real world problems.

There were 5 scientific sessions and 30 presentations (58 co-authors) delivered.

These contributions were centered around the following items:

- Mathematical models and algorithms of combined processing
- New gravity and geomagnetic field modeling techniques. Parallel computing. Comparative analysis of different methods
- Actual requirements for collecting, handling, manipulating and modeling of potential field data, and for improving its quality, reliability, access and exploitation
- Coordination of European Data projects, preparation of initial key data sets of known quality for which consistent instrumental use is desirable

CODATA participation was emphasized by Jacques Emile Dubois (Past President of CODATA International). He presented the CODATA organisation and its scientific objectives. The role of CODATA was also highlighted by Alexei Gvishiani, Chairman, and Jacques Octave Dubois, Vice Chairman of the CODATA Task Group "Data management and Virtual Laboratories". Each gave a scientific presentations and chaired two sessions.

Special acknowledgement was given to Uwe Schäfer and his staff for the excellent organization of this interesting meeting.

Jacques Octave Dubois
Vice-chair CODATA Task Group
"Data Management and Virtual Laboratories"

CODATA Working Group on Biological Collection Data Access

Report from the 2nd Workshop, Royal Botanic Garden, Sydney, November 7-8, 2001

The group was organised in two subgroups, one focusing on protocol development, and one on the definition of the federation schema (content development).

Subgroup on the definition of the Federation Schema

The subgroup on content definition reported on schema development to date. A draft DTD was developed, partly thanks to support provided by the European Natural History Specimen Information Network (ENHSIN). The DTD is based on the results from the first workshop in Santa Barbara, extended by elements from the BioCISE

information model and the British NBN/Recorder model. The schema will evolve through the work of experts on specific "subschemas" such as botanical names or geography. To do so, the schema will be maintained at the group chair's institute and published on the web in a way that makes it understandable for less "XML experienced" experts, i.e. a group of metadata items will form the element definitions which will be put out for discussion, among them: ElementName, ElementTagName, Name in existing standards, Element Description, etc.

A turnaround of 30 days after Request for Comment was considered to be appropriate. The group agreed on the fol-

continued on page 5



NATIONAL MEMBERS SECTION

This section will appear in future Newsletters. Member countries are asked to contribute a short report on data activities and/or general news within their respective country. In this particular section you will find short articles from four of the CODATA National Members. CODATA Committees should submit their articles to: codata@dial.oleane.com

Canadian National Committee for CODATA

At its annual meeting on 9 November 2001 in Ottawa, the Canadian National Committee for CODATA (CNC/CODATA) was pleased to include five representatives from Scientific Unions. The recently-adopted policy of inviting experts from specific disciplines to join the regular CNC/CODATA Members and Observers has proven mutually beneficial and is highly recommended to other national committees. One very tangible benefit realized at this meeting was, for example, the discovery by one of the guests of an ongoing data project in Canada of great relevance to his activities and of which he had been previously unaware. Another interesting feature of the meeting was that two Members, located some 5000 kilometers and 4 time zones apart on the east and west coasts respectively, participated via teleconference.

In addition to dealing with its ongoing business, CNC/CODATA received presentations from the Scientific Union representatives, reviewed progress on the CODATA 2002 Conference in Montreal, formed a sub-committee to prepare nominations for the forthcoming CODATA elections and assigned rapporteur responsibilities for the 2002 version of the Report on Data Activities in Canada. (Earlier Reports may be found at the CNC/CODATA web site:

[http://www.codata.org/canada/.](http://www.codata.org/canada/))

The Canada Institute for Scientific and Technical Information continues to sponsor the Committee.

G. H. Wood
Chairman, CNC/CODATA



Front row (left to right):
Sharon Smith -IPA-,
Mary Zborowski
-Assistant Executive Secretary-
Hanna Dabkowska -Member-

Second row (left to right):
Glen Newton -Observer-,
John Ripmeester -Data on Natural
Gas Hydrates Task Group-,
James Sangster -Observer-, John
Morton -IUPAC-,
Gordon Wood -Chairman-, Alvin
Shrier -IUPS-
Michel Sabourin -Member-

Third row (left to right):
Jean Paul Lauzon -Member-, Andrew
Hakin -Member-,
Paul Mezey -Secretary General-, John
Rodgers -Observer-,
David Lockwood -IUPAP-
Alex Jablonski -COSPAR-

Absent: George Needler -Member-, Francis Ouelette -Member-, Fred Gault -Observer-, Larry Speers -Observer- and Barry M. Wood -Observer-

lowing rules for the future development of the schema: Strong typing will be used. Short element names will be used. Syntactically identical entities will be represented with a single element definition. Namespaces of other communities will be used as far as possible.

continued on page 7



German National Committee for CODATA

The German National Committee held a one-day meeting at the Deutsche Forschungsgemeinschaft (DFG) in Bonn on November 30, 2001. It commenced with a long discussion of the general future politics concerning CODATA and function of the committee itself as well as its membership. These points are presently also a matter of discussion inside the DFG. The next major topic was the review of possible German proposals for the elections of officers and executive committee members of CODATA. Other items discussed were:

- Possible new task groups involving German scientists,
- Suggestions for German participation in the Working Group "Archiving of Scientific Data",
- 18th International CODATA Conference in Montreal (German contributions) and prospective candidates for the CODATA Prize were discussed

A major discussion at the meeting was the report of Dr. Lautenschlager on the progress of a special project on "Citation of Scientific Data". This project was established in autumn of last year by a subgroup of the German National Committee together with other groups of the DFG. It will try to give an answer to the problem that the access to primary scientific data is not guaranteed today because simply these data are very often neither published nor preserved. Stand-alone data publications are broadly not available because no corresponding procedures exist and no personal credits are given to a data publication. The subgroup plans to make a proposal for a solution within Germany towards the middle of this year. It was also decided that the matter should be referred to the next CODATA General Assembly asking for the application of a corresponding new CODATA task group.

Finally, liaisons with other German scientific associations and other committees within the DFG as well as a possible publication on the committee in a German journal were discussed.

H. Behrens
Chair National German Committee

U.S. National Committee for CODATA

The USNC-CODATA is currently planning several exciting international data management and policy initiatives. Other information about the committee's activities may be found on its home page at <http://www.nationalacademies.org/usnc-codata/>

October 2001 USNC-CODATA business meeting.
The committee held its regularly scheduled business meeting on 26 October 2001 to review and plan its national and international activities. The major projects that the Committee is currently working on are described below. The Committee also discussed the "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies," issued in September by the White House Office of Management and Budget (OMB). Ms. Brooke Dickson of the OMB provided an overview of the Guidelines and the Committee discussed issues related to their implementation by the federal science agencies.

International CODATA Working Group on Archiving Scientific Data.

The USNC-CODATA is supporting a U.S. subgroup of this international CODATA working group. The chair of the U.S. subgroup and co-chair of the CODATA working group is Dr. William Anderson, software architect and consultant. The U.S. subgroup met in October 2001 and January 2002 to work on several CODATA working group projects and to discuss plans for future activities. These include: participating in the planning of a data archiving workshop 20-21 May 2002 in Pretoria in collaboration with the South African National Committee for CODATA and the National Research Foundation; (see article pg. 3) the drafting of a report for the 2002 CODATA Conference; organizing a roundtable session for the CODATA 2002 Conference; drafting a proposal for a new CODATA Task Group on the Preservation and Archiving of S&T Data in Developing Countries; and participation in various other related scientific data archiving meetings over the next year. Dr. Anderson gave a presentation on these activities at a "Seminar on Digital Preservation of the

BRAIN TEASER

**“How I need a drink, alcoholic of course,
after the heavy chapters involving quantum mechanics”**

Find the pointless circular argument in this sentence. If you succeed and if you attend the 18th International CODATA Conference in Montreal, you will be presented with one.

Please reply to the secretariat at codata@dial.oleane.com



Record of Science-State of the Art," that is being organized by the International Council for Scientific and Technical Information on 14-15 February 2002, and will be held at the UNESCO Headquarters in Paris.

Workshop on Scientific Data for Decision Making Toward Sustainable Development: Senegal River Basin Case Study.

Scientific databases relating to the environment, natural resources, and health on the African continent are, with few exceptions, very difficult to create and manage effectively. Yet the creation of these and other types of databases—and their subsequent use to produce new information and knowledge for decision makers—is essential to advancing scientific and technical progress in that region and to its sustainable development. This workshop was held in Dakar, Senegal, 13-15 March 2002. The meeting was organized jointly by the USNC-CODATA and the Senegalese National Committee for CODATA. The workshop used a case study of scientific databases that have been created about the Senegal River basin and used for decision making for that region. More detailed information about this project may be found on the Senegal and USNC-CODATA Web sites.

First Inter-American Symposium on Selected Scientific Data Activities.

The USNC-CODATA is organizing this symposium, which will be held 28-29 September 2002 in Montreal, Canada immediately prior to the CODATA 2002 Conference and General Assembly. The purpose of this meeting is to provide a forum for Latin American scientists, data managers, and representatives of science funding organizations to meet with their U.S. and Canadian counterparts to discuss potential activities of mutual interest, exchange information about ongoing and planned research and related data activities, and identify and initiate specific new projects related to data management and exchange. The Symposium will focus on data activities in the environmental sciences and in materials science. The meeting will include an additional focus on data archiving and preservation issues in these two areas, in order to link it more closely to the data archiving activities described above. A follow-up meeting is anticipated to be held the following year in a Latin American country.

2002 CODATA Conference.

The USNC-CODATA is co-hosting with the Canadian National CODATA Committee and international CODATA the CODATA 2002 Conference in Montreal, Canada on 29 September – 3 October 2002. See article pg. 2 in this Newsletter.

Proceedings of the 2000 USNC-CODATA Conference on Data for Science and Society were published in 2001 on the National Academy Press Web site at <http://www.nap.edu/>.

Paul F. Uhler
Director, USNC-CODATA,

47th CODATA Executive Committee Meeting

The Royal Swedish Academy of Sciences has kindly agreed to host the CODATA 47th Executive Committee Meeting at its Academy in Stockholm.

This takes place on 7th and 8th April 2002. It will be followed immediately by a one day meeting on Scientific Data Issues-The Nordic Perspective on 9th of April.

The purpose of this one-day meeting is that CODATA and data experts from Nordic countries can explore issues such as data ownership (IPR), instant dissemination via the web, quality and long term availability of data, with a special focus on the needs of the Nordic Countries.

This is the first time in many years that the Executive Committee meeting has moved from the Paris Secretariat. This may be the establishment of a future precedent!

continued from page 5

Subgroup on Protocol Development

The development of DIGIR (Distributed Generic Information Retrieval), an open source reference implementation <http://digir.sourceforge.net/> was introduced to the group. A design objective in the current work is to decouple the protocol, software, and semantics. One benefit of this decoupling would be to make it easier to evolve or version the federation schema. As long as a stable schema for collection data is not available, the system will use the Darwin Core Version 2 as a simple example. If providers of collection data register their service on a UDDI server <http://www.uddi.org/>, they will not have to inform each portal individually. Portals can poll the UDDI server periodically to discover new providers. Protocol compliant queries will be transmitted to data providers as XML documents. Result sets will also be returned to portals as XML documents. The reference architecture will probably require each provider to publish a "meta data" data set (i.e., collection level data) that will help intelligent portals determine which providers need to be queried to answer a particular user request. The content of meta data items will be discussed on the basis of results of the BioCASE project <http://www.biocase.org>

Walter Berendsohn,
Chair, Biological Collection Data Access
Working Group



IMPORTANT DATES IN THE CODATA CALENDAR

7-8 April 2002

47th CODATA Executive Committee Meeting,
Royal Swedish Academy of Sciences

1 May 2002

Deadline for Call for Papers, 18th International CODATA Conference, 2002

20-21 May 2002

1st CODATA Workshop on Archiving of Scientific and Technical (S&T) Data National Research Foundation, Pretoria

30 April 2002

Deadline for Call for Papers, CODATA Workshop on Information Visualization and Presentation Museum National d'Histoire Naturelle, Paris

5-6 July 2002

13th DSAO Task Group Workshop KISTI, Korea

10-12 July 2002

CODATA Workshop on Information Visualization and Presentation Museum National d'Histoire Naturelle, Paris

6-8 August 2002

1st CODATA Brazilian Conference on Data and Information in Science and Technology, Instituto Nacional de Tecnologia, Rio de Janeiro

28-29 September 2002

First Inter-American Symposium of Selected Scientific Data Activities Montreal

29 September - 3 October 2002

18th International CODATA Conference Montreal

4-5 October 2002

23rd CODATA General Assembly Montreal

Polish CODATA Committee Celebrates Professor Andrzej Maria Bylicki 85th Birthday



The Polish National Committee of CODATA in co-operation with CODATA celebrated the 85th birthday of **Professor Andrzej Maria Bylicki, Ph. D**

with a congratulatory note at its National Committee meeting in November 2001. Professor Bylicki participated in the first CODATA meeting in 1966 and on returning to Poland he organized the Polish CODATA Committee and was appointed Chairman. Born on 17 September, 1916, in Cracow (Poland), he studied in 1934-39 at the Technical University of Lvov, and after the World War II he continued studies at the Jagiellonian University in Cracow, where in 1951 he got his M.Sc. degree in chemistry. He started his professional activity in 1951 at the Institute of Industrial Chemistry in Warsaw as a junior research worker. Since 1956 Professor A.M. Bylicki is a scientific worker of the Institute of Physical Chemistry of the Polish Academy of Sciences in Warsaw. He has got there the doctoral degree in 1957, and he was nominated associate professor in 1965, and full professor in 1973. In 1964-86 he headed the Laboratory of Physical Chemistry of Basic Organic Raw Materials, and in the years 1965-68, 1972-73, and 1977-80 he was a Deputy Director for Scientific Research of the Institute. Since 1985 Professor A.M. Bylicki also heads the Laboratory of Petro-chemistry and Carbo-chemistry of the Polish Academy of Sciences in Gliwice (High Silesia). Professor A.M. Bylicki has published about 70 original papers in the domains of physico-chemical properties of solutions, thermodynamics of phase equilibria, cryometry, properties of coal tar (as a source of organic compounds), kinetics of decarboxylation, thermodynamical data for technology, vapour-liquid and solid-liquid equilibrium, etc. His works were published in the "*Bulletin of the Polish Academy of Sciences, Series of Chemical Sciences*", "*Thermochimica Acta*" "*Journal of Chemical Thermodynamics*", as well as in the "*Proceedings of 6th Int. CODATA Conference*". Professor A.M. Bylicki many times was nominated an official Polish representative at scientific international organs: CODATA, IUPAC, CEPED, etc. He is also a member of the Polish Chemical Society and of Polish Society of Calorimetry and Thermal Analysis. For his scientific activity Professor A.M. Bylicki was rewarded in 1953 by National Scientific Reward of 2nd degree and he also was decorated by high-level national orders.

Professor Juliusz L. Kulikowski
Chairman, CODATA Polish Committee



Editor : Kathleen Cass
CODATA Secretariat
51, bld de Montmorency
75016 Paris

Tél. : 33 1 45 25 04 96
Fax : 33 1 42 88 14 66
e.mail : codata@dial.oleane.com
<http://www.codata.org>