

# CODATA

## Newsletter 96

Committee on Data for Science and Technology

### **CODATA Welcomes Australia As Its Newest Member**

The Australian Academy of Science, represented by its newly formed National Committee for Data for Science, has been accepted into CODATA membership. Realizing that scientific research is increasingly dependent on large quantities of data, such as images of the early Universe, sequences of genes, or surveys of minerals deep beneath the ground, the Academy has taken this strategic step of joining CODATA.

**July 2008**

### **Highlights**

- 1, 2** CODATA Welcomes Australia As Its Newest Member
- 2** The CODATA Task Group for Exchangeable Materials Data Representation...
- 2** CODATA Task Group on Data Sources for Sustainable Development in SADC Countries
- 3** LogKOW® and Mycotox: Canadian Datasets Freely Available Online
- 3** Strategies for Open and Permanent Access to Scientific Information in Latin America...
- 3, 4** CODATA Task Group, Data Sources in Asian and Oceanic Countries (DSAO)...
- 4** Upcoming Conferences
- 4** New Indian National Committee Members
- 4** eGY: Earth and Space Science Informatics Summit, Rome, March 2008

The Academy clearly recognizes that for Australian science to maintain its position in the world and continue to generate new discoveries, it is essential that Australian scientists are able to manage and analyse these data using the best techniques available.



*Dr. Ray Norris*

Infrastructure to enable this is being provided by the Australian Government under the National Collaborative Research Infrastructure Strategy. However, a collaborative approach is also required between scientific disciplines to help tackle the problem. The National Committee aims to promote and facilitate data science across all disciplines of science and provide a national data science voice which can represent Australia at international forums. To achieve this, the Committee will hold regular workshops, promote the development of data management policies and protocols, and promote the adoption of standards for data exchange.

The chair of the new Committee, Dr. Ray Norris of CSIRO, said: 'This is an exciting initiative from the Academy, which will help maintain Australia's position in the front rank of international science...we are now routinely generating terabytes of data, which pose enormous challenges to the way we handle and analyse data. Astronomy, earth sciences, environmental sciences, and life sciences are facing the same challenges, and I'm looking forward to swapping ideas on how to tackle them.'

During 2006, the Prime Minister's Science, Engineering and Innovation Council (PMSEIC) established a Data for Science working group who proposed a comprehensive agenda to ensure that Australia has a systemic approach to enable the best use of data for science. Their report identified a number of challenges currently facing Australian data management, including 'missed opportunities to collaborate' and a 'lack of global engagement'. The new Committee and Australian membership of CODATA will directly address these issues.

**The Committee on Data for Science and Technology (CODATA) was established in 1966 by the International Council for Science (ICSU). The mission of CODATA is to strengthen international science for the benefit of society by promoting improved scientific and technical data management and use.**

<http://www.codata.org>  
<http://dsj.codataweb.org>

*(continued on page 2)*

## CODATA Welcomes Australia As Its Newest Member (continued from page 1)

Members of the National Committee for Data for Science are:

- (Committee Chair) Dr Ray Norris, CSIRO Australia Telescope National Facility
- Ms Kim Finney, Australian Antarctic Data Centre Australian Antarctic Division, Department of Environment and Heritage
- Dr Rhys Francis, Australian eResearch Infrastructure Council
- Dr Alex Held, CSIRO Division of Marine and Atmospheric Research
- Professor Jane Hunter, School of ITEE, University of Queensland
- Dr Tim Littlejohn, Healthcare and Life Sciences, IBM Australia
- Mrs Karen Wilson, National Herbarium of NSW Royal Botanic Gardens
- Dr Lesley Wyborn, Information Services Branch, Geoscience Australia

For more information on the Committee, please go to <http://www.codata.org/membership/nationalmembers/australia.html>

## The CODATA Task Group for Exchangeable Materials Data Representation to Support Scientific Research & Education - meets in New Delhi, India

The CODATA Task Group for Exchangeable Materials Data Representation to Support Scientific Research and Education organized a meeting on 4-5 March in New Delhi, India. It was hosted by the National Physical Laboratory and the Indian National Science Academy. Eight countries participated: Czechoslovakia, France, Netherlands, India, Japan, Russia, South Korea and United States.



The meeting focused on four session topics: 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

A detailed Task Group Meeting report with descriptions of all presentations will be covered in the CODATA *Data Science Journal*. More information on the Task Group can be found on <http://www.codata.org/taskgroups/TGmatisdata/index.html>.

## CODATA Task Group on Data Sources for Sustainable Development in South African Development Community Countries

This first meeting of the TG, to map its project activities, was held in conjunction with the 1<sup>st</sup> African Digital and Information Curation Conference and Workshop in Pretoria on 12-14 February 2008. The presentations from this successful meeting are available at <http://stardata.nrf.ac.za/nadicc/programme.html>.



Two Task Group members presented papers at the conference:

- Dr Alfonse Dubi- "Oceanographic Data and Information Initiatives in Africa".
- Mr Paul Uhlir - "Information Gulags, Intellectual Straightjackets, and Memory Holes";

In addition, Dr Daisy Selematsela facilitated a workshop session in the Social Science and Humanities with focus on "the preservation and sharing of African research data and related documentation".

The Task Group members expressed the value of the conference with regard to their proposed activities.



## **LogKOW<sup>©</sup> and Mycotox: Canadian Datasets Freely Available Online**

The Canadian National Committee for CODATA (CNC/CODATA) regularly seeks to preserve significant Canadian datasets, especially in cases where they are among the best in the world, and where making them accessible to all reduces their risk of being lost or forgotten. CNC/CODATA is pleased to announce two successes:

**LogKOW<sup>©</sup>:** This databank of water-octanol partition coefficients is widely regarded as the best collection of these data in the world. The file contains experimental data, retrieved from the literature, on over 20,000 organic compounds. Although its existence was first announced to CODATA in Newsletter no. 70 (Feb. 1995), it has been available from the CNC/CODATA site since 2003 as a dynamic application, which allows browsing and searching (by either Formula or CAS). Besides partition coefficients, other stored information includes IUPAC names, CAS Registry Numbers, pKa values, S.M.I.L.E.S. notation, experimental conditions of measurement (summarized), recommended Log Kow, and a full bibliography in Chemical Abstracts style. This work of Dr. James Sangster (Sangster Research Laboratories, Montreal, Canada) is available at <http://logkow.cisti.nrc.ca>. (Dr. Sangster is the sponsor of CNC/CODATA's successful [Sangster Award](#) for young scientists in Canada.)

**Mycotox:** The Mycotox file contains a list of fungi, the mycotoxins (toxic metabolites, or metabolites closely related chemically to such toxins) they have been found to produce, and citations to the articles which describes this finding. It is intended for use by mycologists and chemists and assumes a working knowledge of mycology and/or natural product chemistry by its users. The file has been made available by Dr. Alan Taylor on the CNC/CODATA web site since 2007 and is updated annually: <http://www.codata.org/canada/mycotox/>. Inquiries, or offers to collaborate on this project, may be directed to Dr. Taylor through CNC/CODATA.

The mounting of these datasets has been made possible through the ongoing support of the National Research Council's Canada Institute for Scientific and Technical Information (NRC-CISTI). Both of these datasets are retrievable by a search on NRC-CISTI's [online catalogue](#). NRC-CISTI is Canada's national science library and scientific publisher, and is actively addressing the preservation and management of scientific data sets in its current strategic activities.

By Mary Zborowski, Executive Secretary, Canadian National Committee for CODATA

## **Strategies for Open and Permanent Access to Scientific Information in Latin America: Focus on Health and Environmental Information for Sustainable Development**

The report on the Workshop on: *Strategies for Open and Permanent Access to Scientific Information in Latin America: Focus on Health and Environmental Information for Sustainable Development*, which took place in Sao Paulo, Brazil can be accessed on [http://www.cria.org.br/eventos/codata2007/report/Brazil2007\\_workshop\\_report.rtf](http://www.cria.org.br/eventos/codata2007/report/Brazil2007_workshop_report.rtf)

## **CODATA Task Group, Data Sources in Asian and Oceanic Countries (DSAO), meet around a Joint conference organized by it and the Chinese National CODATA Committee, UN GAID and IAP**

Over 80 participants attended a Joint Conference organized by the Chinese National CODATA Committee, the United Nations Global Alliance for ICT and Development (U.N. GAID) e-SDDC and the InterAcademy Panel (IAP) on International Issues. The Conference was hosted by the Computer Network Information Centre, the Chinese Academy of Sciences (CAS). It attracted speakers from Beijing, Taipei, Hong Kong, Malaysia, United States, and Sweden. The meeting was opened by CAO Jinghua, Deputy Director General of the Bureau of International Co-operation of the Chinese Academy of Sciences. HUANG Xiangyang, (Director General of the Computer Network Information Centre of the Chinese Academy of Science) presented a talk in the Plenary Session along with GUO Huadong, (Deputy Secretary General of Chinese Academy of Sciences), Paul Uhlir (US National Academies), Lars Hernroth (Royal Swedish Academy of Sciences), YAN Baoping and LIU Chuang (Chinese Academy of Sciences). The Task Group also met on December 5<sup>th</sup> to discuss its future activities.

### **Joint Conference by CODATA TGs, IAP and UN GAID e-SDDC**



(continued on page 4)

## DSAÖ (continued from page 3)

An International Planning Meeting of the InterAcademy Panel on International Issues (IAP) on *Promoting Access to and Use of Digital Knowledge Resources in Countries with Developing and Transitional Economies: Focus on the Role of Science Academies in Asia-Pacific Region* was held aiming to develop an Action Agenda for Science Academies in Promoting Access to and Use of Research and Education Networks (RENs) in the Asia-Pacific Region, as well as to develop an Action Agenda for Science Academies in Promoting Access to Digital Knowledge Resources (including open Institutional Repositories) in the Asia-Pacific Region.

The UN GAID e-SDDC workshop titled “*The Global Alliance for Enhancing Access to and Application of Scientific Data in Developing Countries (e-SDDC)*” of the United Nations Global Alliance for ICT and Development (UN GAID): Building a decentralized network of networks (Networking)” was also successfully held. The aim of the workshop was to discuss the implementation strategy for the e-SDDC activity .

To find out more details on speakers and presentations see <http://dsao.codata.cn/> . ☐

## Upcoming Conferences

**Environmental Informatics and Industrial Ecology**  
 22nd International Conference on Informatics for Environmental Protection (EnviroInfo 2008)  
 10-12 September 2008  
 Lüneburg, Germany  
<http://enviroinfo2008.org> ☐

## New Indian National Committee Members

The Indian National Committee for CODATA has been reconstituted as of January 1, 2008. The members of this new committee, effective for a period of four years, are:

*Chairman:*

Dr.Baldev Raj, FNA  
 Distinguished Scientist & Director  
 Indira Gandhi Centre for Atomic Research

*Members:*

Dr.V.P.Dimri, FNA, Director  
 National Geophysical Research Institute

Professor Alok Bhattacharya, FNA  
 School of Life Sciences, Jawaharlal Nehru University

Dr.K.Sekar, Principal Research Scientist  
 Bioinformatics Centre, Indian Institute of Science

Professor Jitendra Kumar, Professor & Head  
 Materials Science Programme, Indian Institute of Technology ☐

## eGY: Earth and Space Science Informatics Summit, Rome, March 2008

The Summit took place at the Italian Geographical Society and home of the International Geographical Union, Rome on 13 and 14 March 2008. The main theme of the summit was *geoinformatics aspects of eGY and developments worldwide in Earth and space science informatics*. Earth & Space Science informatics bodies and programs are expanding rapidly around the world in response to three drivers (i) modern ICT capabilities, (ii) demand for ready access to data and services in order to tackle modern cross-disciplinary, complex problems as well as traditional ones, (iii) greater readiness by governments to support a science information commons. One of the objectives of the Summit was to help to produce a unified approach to common issues, and reduce replication of effort. The CODATA Working Group *eGY Earth and Space Science Data Interoperability*, actively participated at the event. Alexei Gvishiani, Chair of the Group, reported on its activities.



The slides of the report are available at <http://egy.gcras.ru>. The meeting was attended by leading representatives of ICSU, ICSU-SCID, IUGG, IUGS, IGU, IAGA, IASA, EGU, AGU, WDC Panel, GSA, NCAR, ESA, ESRIN, GEOSS, WGISS, USGS, RAS, Australia, China, and Italy. Participants at the Summit represented the interests of more than 40 organisations. A full report about the Summit is on the eGY website: <http://www.egy.org/>. ☐

### CODATA Newsletter

Editor: Gordon Wood  
 Layout: Cheryl Williams Levey

CODATA Secretariat  
 5 rue Auguste Vacquerie  
 75016 Paris, France  
 Tel: +33 1 45250496  
 Fax: +33 1 42881466

Email: [codata@dial.oleane.com](mailto:codata@dial.oleane.com)  
 Website: <http://www.codata.org>