http://www.egy.org/

*e*GY News No. 3 December 2004

Editor's Note

December's holidays took their toll – I have not received a few expected contributions to the No. 3 issue, and now I am facing a dilemma – either to cancel the issue or still provide you some information. I have a feeling that if the next couple of issues will be ending this way, then we should think stopping publishing the *e*GY News on a regular basis. The community newsletter cannot be supported only by a few people; without the external contributions the *e*GY News will be dead. Thus, everybody who is interested in keeping this News alive should think of some contribution… even it would be a list of various geoscience data sites or virtual observatories (or similar efforts) in different disciplines.

Fortunately to us, the ICSU has recently produced its long awaited assessment of scientific data and their management: http://www.icsu.org/Gestion/img/ICSU_DOC_DOWNLOAD/551_DD_FILE_PAA_Data_and_Information.pdf). The 43 pages long report provides an excellent plank for the entire *e*GY initiative. In this issue we abstracted only key elements of its Executive Summary (5 pages) focusing on the most relevant to *e*GY recommendations. We strongly encourage every scientist and scientific administrator to read this report.

The essence of *e*GY

"What is eGY?" This is the first question that people ask. The simple answer is that it is an opportunity. An opportunity for the geoscience community to use the 50-year anniversary of the IGY in 2007-2008 for an international effort to boost *e*-Science for Geoscience. The challenges are to improve data discovery, data release, data access, processing ability, data integration and knowledge discovery, and data preservation and undertake capacity building and outreach programs. *e*GY as an organization provides the coordinating environment and stimulus to encourage this to happen. The results that will be achieved will depend on how much we as a community take advantage of this opportunity.

(Received from Charlie Barton)

Scientific Data and Information - Report of the CSPR Assessment Panel, ICSU, December 2004

A report of the CSPR Assessment Panel

Execute Summary Excerpts

Page 7:

The ICSU Priority Area Assessment (PAA) Panel on Scientific Data and Information strongly recommends that ICSU assume an international leadership role in identifying and addressing critical policy and management issues related to scientific data and information.

The Panel recommends that, on the basis of the many specific recommendations in this report, ICSU develop a long-term strategic framework for scientific data and information (policies, practices and infrastructure). An essential part of the development of this framework should be the closer coordination, and in some instances the transformation, of ICSU's current data and information activities.

The panel recommends that, in parallel to the development of the long-term strategic framework, ICSU establish an international Scientific Data and Information Forum (SciDIF) involving all the key stakeholders: ICSU members, interdisciplinary bodies, science funding bodies and other data providers and users. Through SciDIF, ICSU should aim to ensure that the full benefits of new data and information technologies and capabilities are extended to scientists throughout the world.

Recognizing the importance and urgency of these issues, ICSU should establish a Strategic Data and Information Committee to oversee the development of a long-term integrated framework for data and information and a Scientific Data and Information Forum (SciDIF). Membership of this committee should include representatives of relevant ICSU bodies and unions, and experts in information technology and professional data management. This should be an *ad hoc*

committee, with members (10-12 max.) appointed for a period of three years.

Page 8:

DATA AND INFORMATION RESCUE

Many types of data are at risk. They are not being used for scientific research because they are not available in digital formats or they are in danger of being lost because the media on which they are recorded may decay, become corrupted, or be superseded by new software. This is a particularly acute problem in developing and transitional countries.

9. Scientists should inventory major collections of extant data and information and should set priorities for the rescue and permanent preservation of the data and information that are most valuable and at greatest risk.

10. ICSU and its members should draw the attention of scientists, public policy makers, and research foundations to the issue of data at risk and ways to deal with the issue.

11. Scientists or countries that undertake significant efforts in the rescue of data at risk should consider the advantages of public/private partnerships in this effort.

Page 9:

PROFESSIONAL DATA AND INFORMATION MANAGEMENT

Scientific data and information management can no longer be viewed as a task for untrained amateurs or as part of the routine "clean up" at the completion of a research project. The use of advanced information technology in scientific data management and dissemination makes it essential that data management be the responsibility of professionals. Scientific data centers and archives require stability in their financial resources so that they can make institutional commitments to data management and preservation over many decades.

16. The panel recommends that ICSU play a major role in promoting professional data management and that it foster greater attention to consistency, quality, permanent preservation of the scientific data record, and the use of common data management standards throughout the global scientific community.

17. Recognizing that scientific data and information management is undergoing rapid innovation and change, information technology specialists, librarians, research scientists, government data producers, donors, and others should be involved in a concerted effort to develop standards and curricula for professional training for scientific data managers.

18. Financial support for data and information management should become a routine component in all research budgets and the evaluation criteria for assessing research funding proposals should include evaluation of data management.

19. All scientists should receive training in data management as part of their graduate and postgraduate education. ICSU should encourage the development of guidelines for data management by working scientists and their institutions.

20. Scientists should be recognized and given credit for the scientific contribution of the data sets that they produce as well as for the analysis of those data.

21. ICSU, its members and associated bodies should raise awareness of the increasingly important role that institutional repositories play in relation to scientific information management and preservation and the need to ensure that such repositories are properly resourced, developed and maintained.

ARCHIVING

Permanent archiving of scientific data and information is essential. In some fields, there are institutions that have a clear responsibility for data archiving, but this is not always the case and varies from one scientific field to another. There is a distinction between data centers, which are responsible for providing immediate access to scientific data and information, and archives, which provide for permanent preservation and management of data and information.

26. ICSU and its members should raise awareness of the need for long-term institutional support for data archives both at the national and international level.

27. ISCU should foster discussion within the scientific community, including its members and interdisciplinary bodies, on criteria, institutional structures, and models for decision making related to the permanent preservation of scientific data and information.

28. Ways to reduce the costs of archiving, such as sampling from extant data bases or establishing multiple classifications to prioritize levels of archival support, should be examined by the scientific community.

Page 10:

EQUITABLE ACCESS TO DATA AND INFORMATION

Science has long been best served by a system of minimal constraints on the availability of data and information. A strong public domain for scientific data and information promotes greater return from investment in research. It stimulates innovation and enables more informed decision-making.

33. ICSU should continue to actively promote the principle of full and open access to scientific data.

34. With regard to scientific publishing, ICSU should ensure that the principle of universal and equitable access to scientific publications is upheld.

THE DIGITAL DIVIDE

The digital divide is most evident in low- and medium income countries, but it has an impact on scientists throughout

the world because it limits potential research and access to data and information in many regions.

35. Key issues relating to the digital divide in scientific data and information are addressed in the Science in the Information Society Agenda for Action. The panel strongly endorses the actions relating to the provision of high-speed internet access, human capacity building, electronic publishing and initiatives to make scientific information more accessible. ICSU should work with its members and partners to encourage the implementation of these actions.

36. Emphasis should be placed on the need for professional data and information management in those countries where the scientific data infrastructure is being constructed. ICSU bodies, such as the International Network for Access to Scientific Publications (INASP) and the Committee on Data for Science and Technology (CODATA), have a key education and training role to play in these countries.

SYSTEMS FOR DATA DISSEMINATION

The ease of disseminating data via the Internet encourages individual scientists to make their research data available to others through personal or project web pages or through distributed federations of data providers. However, there is concern that data managed by individuals or research groups in voluntary distributed systems on the Internet may not continue to be available over time.

37. ICSU should bring together representatives of voluntary data confederations and distributed data systems to discuss what has been learned over the past ten years about what contributes to the success of voluntary data confederations, what undermines them, and what must be done to preserve and enhance access to the data in the future.

WHO PAYS?

Data production and management are costly. Collection of data, preparation of metadata, and provision of professional data management expertise and institutional support for data dissemination and permanent archiving will add to the overall expense of specific research projects and maintaining the larger research infrastructure. There are a number of economic models for supporting the necessary scientific data and information management, but none of them is completely satisfactory.

38. Ensuring the long-term accessibility of increasing quantities of scientific data and information will necessitate increased public (and private) investment in data management and long-term institutional support. ICSU and its members should explore various solutions to meet the financial challenge of providing full and open access to scientific data and universal and equitable access to publications.

Page 11:

REVIEW OF ICSU BODIES

COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY (CODATA)

In the context of the development of a long-term strategic framework and international Scientific Data and Information Forum:

48. CODATA should develop a clear long-term strategy that focuses on key international data management and policy issues, giving special attention to the needs of developing countries.

52. ICSU should play a stronger role in the Global Observing Systems by fostering the development and implementation of appropriate policies and data management procedures and representing scientific data user needs. There is a need for both operational archives for dissemination of data and a strategy for long-term data preservation.

53. ICSU should be actively involved in the Group on Earth Observations (GEO) as a representative of the international scientific community, advocating appropriate policies for scientific consultation, data collection, data access, and professional management of Earth observation data.

PANEL ON WORLD DATA CENTERS (WDC)

55. ICSU should re-examine its entire data center infrastructure in light of technological and scientific changes in data collection, use, and management. Planning for an updated system should build upon the successful accomplishments of the World Data Centers, but should go beyond current practice to take advantages of new technologies and capabilities. This effort should be integrated into the development of a long-term strategic framework and international Scientific Data and Information Forum (SciDIF). It must take account of the needs of existing ICSU programs and other new initiatives, such as the Group on Earth Observations (GEO), the International Polar Year 2007- 2008, and the Electronic Geophysical Year (*e*GY).

Forthcoming eGY Meetings

February 2005: *e*GY Symposium and Workshop (Boulder).

23-27 May 2005: Spring AGU/SEG/NABS/SPD/AAS; New Orleans, USA; U08 *e*GY: e-Science for Geoscience.

18-29 July 2005: IAGA Scientific Assembly, Toulouse – *e*GY special session.

Editor: Volodya Papitashvili, papita@umich.edu