

CODATA

Newsletter 90

Committee on Data for Science and Technology

Priority Area Assessment Report on Data and Information: Interview with Roberta Balstad

In December 2004, The International Council for Science, ICSU, published its Priority Area Assessment Report on Data and Information. Some of the recommendations contained in this 42-page report focused directly and/or indirectly on CODATA. In order to learn more about some of the issues raised in the report and the role of CODATA in the implementation of these recommendations, Kathleen Cass, CODATA Executive Director, interviewed the Chair of the ICSU Panel responsible for the report, Roberta Balstad.

April 2005

Highlights

- 1,2,3** Priority Area Assessment Report on Data and Information: Interview with Roberta Balstad
- 4** Anthropometry, A Big Word With Human Dimensions
- 4** Upcoming Events
- 4** In Memoriam, Prof. Gabriel Olaere Ajayi
- 4** CODATA 2004 Proceedings Now Available

<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 inter-disciplinary basis, CODATA seeks to improve the quality, reliability, management and accessibility of data of importance in all fields of science and technology.

Roberta, on behalf of CODATA, I would like to congratulate you and the committee on this report which contains what many in the CODATA family have described as, "a comprehensive and visionary review" of data and information issues. I would also like to thank you for taking the time to address some of the questions that have been raised within the data community since the report was published.



Roberta Balstad

Q. How has the report been received to date by the scientific community and are other organizations embracing it?

I think we have all been pleased at the positive response to the report in the scientific community. There are scientific organizations that have already begun to address the issues raised in the report, such as CODATA and the World Data Centers, and others that are beginning to discuss the issues that it raises within and across disciplines.

However, the PAA panel recognized from the start that issues relating to scientific data and information are no longer solely under the control of scientists. Many financial and policy decisions that affect scientific data and information collection and dissemination are now made by national governments or private corporations, and others are made by funding agencies and foundations.

For this reason, the PAA report was intended to speak to a broad audience. The primary audience, of course, consists of ICSU, members of the ICSU family, and the scientific community, but the panel also hoped that ICSU and the scientific community would communicate the PAA recommendations to national and multinational governments, to organizations that provide financial support for research, and to both for-profit and not-for-profit groups. The collection and preservation of scientific data and information is essential to the current and future health of science, and ensuring that these data will be available to scientists now and in the future requires close collaboration across the scientific and policy communities.

(continued on page 2)

Priority Area Assessment Report on Data and Information: Interview with Roberta Balstad (continued from page 1)

Q. Roberta, in your opinion, would there have been a need for such a report on data and information within ICSU ten years ago?

If yes, do you think the recommendations in the report *vis a vis* ICSU would be similar to those contained in the current report?

If not, what has been the greatest development within the scientific community to merit such a review?

I don't believe that the report *could* have been written ten years ago. Certainly there were many people back then who recognized the need for large scale, multi-user scientific data bases and for disseminating data over the newly invented World Wide Web, and there were others who were experimenting with combining many data bases in meta-analysis, providing scientific information over the Internet, and developing new types of software for data mining and spatial visualization and analysis. But ten years ago, the scientific potential of combining these technologies was just beginning to be recognized and the field of data and information management was just emerging. I think it is safe to say that the conduct of science has been transformed over the past decade through information technology, network access, new data base development, and a scientific research emphasis on the dynamics of change over time.

Q. It is true that data and information issues impact on many scientific disciplines and are therefore interdisciplinary by nature. Is an interdisciplinary organization the most appropriate mechanism to address these issues, or should these issues be left to the individual scientific disciplines to address on an *ad hoc* basis?

The issue of scientific data collection and management has traditionally been addressed on an *ad hoc* basis, first by individual scientists, who generally felt that they had to collect their own research data and often even built their own instruments or sensors. Later, groups of scientists in the same or related disciplines began to collaborate on the development of larger data bases for general use. This approach had the added scientific advantage of encouraging research testing hypotheses on the same body of data. What is new today is the recognition that data cannot be defined solely in disciplinary terms, that they describe phenomena that may be studied from multiple perspectives and in the context of multiple disciplines and interdisciplinary fields. Many of the leading scientific questions being addressed today are themselves multi- or trans-disciplinary, and scientific research draws upon relevant data regardless of the disciplinary affiliation of the individuals who were initially responsible for their collection. Scientific questions in the future are likely to be even more eclectic in terms of data needs.

For these reasons, it is essential that the scientific community develop data standards and management practices across the current disciplinary configurations in academia.

Information management is a different issue. Scientific information, particularly journal articles, is generally reviewed and disseminated within disciplines, often by commercial publishers or by non-for-profit groups that depend upon publication income for other purposes. In part because the dissemination of scientific information is so decentralized and in part because it generates needed income for the enterprise, we have not yet solved the problem of providing non-discriminatory access to scientific information. This is particularly true in terms of providing scientists in the developing world with access to publications, although INASP provides an excellent model for dissemination that could be applied elsewhere.

Q. One of the terms of reference addressed to the panel was to “define an overarching ‘mission’ and role for ICSU in the area of *Scientific Data and Information.*” What is this mission and what role has CODATA in implementing this mission?

The PAA panel believes that ICSU should “assume an international leadership role in identifying and addressing critical policy and management issues related to scientific data and information.” The report emphasizes that ICSU is already involved in scientific data and information through many of the groups in the ICSU family, so the fact of an ICSU role is not new. But ICSU's role needs to expand as large-scale data collection, data sharing, and information technologies have been changing the scientific enterprise over time. Only ICSU, working with the organizations in the ICSU family, the national academies, and the international disciplinary associations, is able to work across the scientific and the policy communities (both national and multinational) and to represent the full range of scientific research.

CODATA will have a major role in these activities. As you know, I am chair of the US National Committee on CODATA and believe that CODATA plays a vital role addressing data and information policy issues in many nations. I would like to see that role expanded into more countries, particularly developing countries where the tradition of scientific data collection, sharing, and management is less well developed. One of the basic messages of the PAA Report is that CODATA needs a clear long-term strategy, while focusing its short term activities on key international issues, such as the World Summit on the Information Society (WSIS). The CODATA strategy should be developed in consultation with ICSU to avoid duplication of responsibilities.

(continued on page 3)

Priority Area Assessment Report on Data and Information: Interview with Roberta Balstad (continued from page 1)

Q. The report recommends that “The lines of communication between CODATA and ICSU need to be improved” and that “CODATA should continue to develop a closer working relationship with ICSU bodies such as INASP and ICSTI in areas where there are complementarities and clear added value.” Do you have any specific suggestions on how CODATA might approach this?

Again, the panel felt that CODATA and ICSU should strategize together to take advantage of each organization’s strengths. There is too much work for any single organization to do alone.

It is my understanding that CODATA has already begun to work in concert with ICSTI and INASP, which is why the panel recommended that CODATA “continue” to develop a closer working relationship with the other two organizations. For example, since INASP works on information access, and CODATA on data access, there could be a natural synergy in the two organizations, particularly if CODATA decides to emphasize its work with developing countries.

Q. The report highlights the need for CODATA to strengthen its membership and recommends that “ICSU should encourage those of its members who are not currently affiliated to CODATA to reconsider this position.” What do you think CODATA could do to strengthen its membership?

This is a matter for widespread discussion within CODATA and should be part of the strategic planning process. CODATA has traditionally had national members and obtains financial support through these members. But it may decide that it is better to seek a different type of member, such as major archives or data organizations. Or it could decide to obtain its financial support through a combination of some national members, some organizational members, and some grants-in-aid. For example, if CODATA were to define its future role primarily in terms of developing countries, it could seek financial support for its activities through national or multilateral science agencies and development agencies. This approach has worked very well for INASP.

Q. In its conclusion the report recommends the need for a long-term strategic framework for scientific data and information and recommends also the establishment of an International Scientific Data and Information Forum (SciDIF). What is the time span for implementing these activities, and what role do you see for CODATA in these efforts?

The basis for developing the strategic framework is provided in the PAA and so to some extent the efforts of CODATA and other bodies in implementing the PAA recommendations are already contributing to its development. However, the recommendation of the Panel is that ICSU should establish an *ad hoc* panel for three years to guide and monitor this process and an International Forum where all the relevant stakeholders from both within and outside the ICSU family can be involved. ICSU does not have the financial resources to induce change but it is uniquely placed to bring together leading scientists from different disciplines and countries and other stakeholders in scientific data and information. All of these need to be persuaded of the importance and urgency of the issues raised in the PAA, if they are to be fully addressed. This will not happen immediately but I am optimistic that over the next 3 years we will see significant progress. A key milestone will be the ICSU General Assembly in China in October this year, when Members will be asked to approve the establishment of the new *ad hoc* strategic committee and International Forum.

CODATA has a critical role to play, firstly in defining its own strategy and priorities, taking account of the PAA recommendations, and secondly by working with the new *ad hoc* committee and Forum to ensure the necessary changes. In some areas, I expect that CODATA will be the lead organisation, in others it will make an active contribution in association with other ICSU bodies or organizations and in others still its role may be more limited. These choices need to be laid out in CODATA’s own strategy for the next few years. We hope that the PAA provides the impetus for CODATA to develop and expand and to play a leading role in international interdisciplinary data issues into the future.

Q. What did you find the most challenging aspect of this review?

The most challenging aspect of the PAA was pulling together so many diverse threads related to data and information, changes in information technologies, and changes in ways of conducting science and doing so in a short report that speaks to scientists, policy makers, and data and information professionals.

Q. And the most rewarding aspect?

The most rewarding aspect is easy—it was working with an excellent PAA panel, meeting the committed and talented people who work in the various organizations in the ICSU family, such as CODATA, and knowing that we are all dealing with issues that are critical for the future of science. ■

Anthropometry, A Big Word With Human Dimensions

CODATA recently approved the establishment of the Anthropometric Data and Engineering Task Group. The study of human body measurement for use in anthropological classification and comparison is anthropometry, according to the American Heritage Dictionary. We've all heard the saying that people come in all shapes and sizes, this group intends to make the data that represents those shapes and sizes accessible and usable for engineering applications.

Generally, people from the Netherlands are tall, from Japan are smaller, and customers of automobiles, baby cribs, aircraft seats and clothes come in all sizes. Manufacturers of these products want to produce goods that are safe and comfortable for as many people, in a target population, as possible.

Imagine trying to get a representative sample population and measuring the shape of that population. These "anthropometric surveys" are expensive, time consuming and fraught with difficulty. In spite of these difficulties, many anthropometric surveys exist and are valuable sources of information. Integrating these data sources, however, is challenging and is also a key goal of this task group. Ideally we will produce a system to allow intelligent access to anthropometric data that is useful to designers and manufacturers of engineered products of all types.

(Contributed by Sandy Ressler) ■

Upcoming Events

ISGI CODATA Symposium, Berlin, Germany, 14-16 September 2005.

<http://www.codata-germany.org/ISGI>

Ensuring Long-term Preservation and Adding Value to Scientific and Technical data (PV 2005), 21- 23 November 2005, Royal Society, Edinburgh, UK

<http://www.ukoln.ac.uk/events/pv-2005/>

19th International Conference Informatics for Environmental Protection Networking Environmental Information, September 7- 9, 2005, Brno, Czech Republic

<http://www.enviroinfo2005.org>

Multimedia. Where Do We Go From Here? CODATA International Symposium, September 19-20, 2005, European Academy, Berlin, Germany

http://www.codata.org/announcements/Berlin_2005_CfPF.htm ■



In Memorium Prof. Gabriel Olaere Ajayi

The Nigerian CODATA and the Nigerian Information Technology community suffered a loss with the passing on of Nigeria's internet Administrative

point of contact, Prof. Gabriel Olaere. Ajayi on 12 December 2004. Prof. Ajayi was the pioneer Director-General and Chief Executive Officer of the National Information Technology Development Agency.

Prof. Ajayi was born on 30 November 1941. He read Physics at the University of Nigeria, Nsukka, graduating with a first class in 1965. He later enrolled for a Master's degree in Electronic and Telecommunications Engineering at the Ahmadu Bello University and graduated in 1968.

He also held an M.Sc in Communications Engineering from the University of Manchester in 1971 and a doctorate (Ph.D) in Electrical Engineering also from Manchester. He joined the University of Ife, Ile-Ife, Nigeria, where he rose to become a Professor. He had many publications to his credit. He was very visible at the CODATA Africa workshop held in Dakar, Senegal and the CODATA conference in Baveno Italy, both in 2000. He was a pillar of support to the Nigerian CODATA.

The CODATA family will miss Prof. Ajayi very much and we wish him eternal rest.

(Contributed by Kingsley Oise Momodu) ■

CODATA 2004 Conference Proceedings Now Available

Thank you for your participation in the 2004 CODATA Conference. The proceedings are now available electronically on the CODATA website at:

<http://www.codata.org/04conf>

If you have not already submitted your manuscript, please do so as soon as possible. Submission guidelines can be reviewed here:

<http://www.codata.org/04conf/publguidelines.html>

Your submission will also be considered for publication in the *Data Science Journal* at

<http://www.datasciencejournal.org>.

CODATA Newsletter Editor: Cheryl Levey

CODATA Secretariat
51, Boulevard de Montmorency, 75016 Paris, France
Tel: +33 1 45250496 / Fax: +33 1 42881466
Email: codata@dial.oleane.com
Website: <http://www.codata.org>