Workshop on Big Data for International Scientific Programmes: Challenges and Opportunities

Beijing, 8-9 June 2014

Summary

CODATA, the ICSU Committee on Data for Science and Technology, will convene a *Workshop on Big Data for International Scientific Programmes: Challenges and Opportunities* to be held in Beijing, China, on 8-9 June 2014.

The workshop is designed to provide a better understanding of the opportunities and challenges of 'Big Data' for international collaborative science programmes, including ICSU-sponsored programmes such as Future Earth and Integrated Research on Disaster Risk (IRDR), as well as international initiatives such as the Group on Earth Observations (GEO), the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES), the International Society for Digital Earth (ISDE), and initiatives sponsored by the Belmont Forum. The results of the workshop will also help set an agenda for CODATA activities on Big Data for international science.

The workshop is co-sponsored by the ICSU World Data System (ICSU-WDS), the Research Data Alliance (RDA), the International Society for Digital Earth (ISDE) and the Institute of Remote Sensing and Digital Earth of the Chinese Academy of Sciences (RADI).

The event will follow the Future Earth Scientific Committee and Engagement Committee Meetings being held in Beijing and take place in parallel with the 2nd IRDR International Conference to facilitate participation by attendees of those communities.

Context and Rationale

Rapid advances in technology are radically changing the way in which data are being collected, used, and stored. Digital data are gathered, replicated, moved, and processed more quickly and in greater volumes than ever before. As new information technologies, sensors, and communication networks develop, the range and complexity of scientific data continue to grow. With data volumes expanding beyond the petabyte and exabyte levels across many scientific disciplines, the capacity for storage and preservation and for long-term use may be exceeded in many fields. Above all, the opportunities to extract information from complex data sources from diverse disciplines offers compelling reasons to embrace the new scientific methods and approaches of 'Big Data' and data-driven research.

In a world dealing with growing populations, pressing economic and social needs, natural and technological hazards, and climate change, there is a clear need for more, robust and high-quality data—along with new analytics and models and faster delivery and visualization of information—to support evidence-based decision making and risk management by a wide range of stakeholders. There are lessons to be learned, both positive and negative, from big data efforts in genomics, business, astronomy, and other fields that can be applied in developing Future Earth, IRDR, and other programmes and in ensuring that these initiatives have greater and more lasting impacts than they might otherwise have.

'Big Data' presents particularly significant challenges and notable opportunities for transdisciplinary, international research programmes such as Future Earth and IRDR, as well as for international initiatives such as GEO, IPBES, and the Belmont Forum. These efforts aim to guide research and produce research results and data in ways that improve decision-making on critical issues for humankind and the environment. To accomplish this requires the integration of increasingly diverse and complex datasets and the extraction and

interpretation of knowledge utilizing sophisticated statistical techniques, complex simulation models, and other computationally intensive approaches.

'Big Data' has become an overarching and critical issue for both basic and applied scientific research, and one to which CODATA brings expertise and leadership. As an interdisciplinary body of ICSU, CODATA has a convening role and a mission to explore data issues in the context of international research, whether specific disciplines (as represented by International Scientific Unions) or the interdisciplinary research as promoted in international, collaborative programmes.

Indicative Workshop Programme

The purpose of the *Workshop on Big Data for International Scientific Programmes* is to bring together Big Data leaders from discipline-specific data efforts, major interdisciplinary programmes, and national and international organizations and activities to help identify opportunities and approaches for improving cooperation and collaboration in applications of Big Data in international scientific programmes and initiatives.

'Using Big Data to Advance International Scientific Programmes', 8 June: Day One of the two-day workshop will focus on exploring and reviewing current Big Data-related programmes, with special emphasis on how communities cooperate and approaches used to overcome issues.

Session 1.2 Data Challenges and Big Data in Future Earth, IRDR, GEO and ISDE

Session 1.1 Big Data Today: lessons from international science initiatives and the commercial world

Sessions 1.3 and 2.1 Success and Lessons Learnt in Data Intensive Science

Reception and Banquet sponsored by RADI ([Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences](#)).

 'Building International Cooperation and Collaboration on Big Data for International Science', 9 June: Day Two will start by identifying opportunities for new levels of cooperation and coordination, especially for emerging large-scale S&T programmes. The workshop will conclude with discussions on the role of CODATA in facilitating Big Data progress in S&T through working groups and other means.

Session 2.2 How to make progress: Opportunities for Collaboration on Big Data for International Science

Session 2.3 Panel to discuss CODATA WG on Big Data for International Science

The programme will feature keynote and other invited speakers from leaders in Big Data. There also are discussion sessions designed to identify key opportunities for new levels of cooperation and coordination among international scientific programmes, as well as contribute to defining the scope of a CODATA activities in Big Data.

Further Information

The workshop is envisaged as a relatively small and focused event, convening 40-50 world experts. It is intended that proceedings and a report of the workshop will be published in the [*CODATA Data Science Journal*](#), while a position piece will be submitted to a high impact journal. Above all, the workshop will bring together Big Data experts and researchers involved in international research programmes and identify opportunities for future collaboration to facilitate specific research objectives. These themes will feed into [SciDataCon 2014, the International Conference on Data Sharing and Integration for Global Sustainability](#) which ICSU CODATA is organizing with its sister organization, the [ICSU World Data System](#). Finally, it is intended that the workshop should help set the agenda for ongoing CODATA activities in support of 'Big Data for International Science'.

The workshop will be supported by the Institute of Remote Sensing and Digital Earth (RADI) of the Chinese Academy of Sciences (CAS), and by other partner organizations. The full programme and other details will appear shortly.