Brief for GSDR – 2016 Update

Open Data in a Big Data World: challenges and opportunities for sustainable development

Charles Ebikeme[1], Simon Hodson[2], Geoffrey Boulton[2], Heide Hackmann[1], Anne-Sophie Stevance[1], Lucilla Spini[1]

[1]International Council for Science (ICSU), Paris, France, [2]ICSU Committee on Data for Science and Technology (CODATA), Paris, France

Introduction

The famed 'digital revolution' has created an unprecedented explosion in the capacity to acquire, store, manipulate and instantaneously transmit vast and complex data volumes, such that the 'digital revolution' has initiated a 'data revolution'. The ability to quantify and measure almost every aspect of life presents an opportunity and a challenge for the research endeavour and for the knowledge built on top of that.

Openness, the lack thereof, and 'intelligent openness' (Royal Society, 2012) is central to obtaining the promise of the data revolution (UNEAG, 2014). Too often, recently, we have seen private parties engaged in the walling-off of knowledge by restricting access to the data underpinnings of that knowledge.

The digital revolution provides a powerful medium through which scientific productivity and creativity can be enhanced by permitting data and ideas to flow openly, rapidly and pervasively through the networked interaction of many minds. If this social revolution in science is to be realised it is vital that we adopt a default position that publicly funded data should be made publicly accessible and re-usable.

The capacity to access and use data is deeply significant and transformative for a range of research fields both in the natural and social sciences: in weather and climate forecasting; in understanding the workings of the brain; in the behaviour of the global economy; in evaluating agricultural productivity; in demographic forecasts; in unravelling histories; and in many of contemporary global challenges such as those of environmental change, infectious disease and mass migration that require combined insights and data from many disciplines.

New institutional paradigms for promoting open data

Recently, four major organisations representing global science, the International Council for Science (ICSU), the InterAcademy Partnership (IAP), The World Academy of Sciences for the advancement of science in developing countries (TWAS) and the International Social Science Council (ISSC), announced a global accord proposing principles and practices for open access to research data.

The accord titled "Open data in a big data world" (ICSU, 2015) takes a coordinated step forward with the bodies that advocate and support open access, and the academies that influence government, policymakers and funders.

The accord identifies the opportunities and challenges of the data revolution as today's *predominant* issue for global science policy. It proposes fundamental principles that should be adopted in responding to this issue. These principles are consistent with ones being carried out in practice in some national research systems and in some disciplinary fields. It adds the distinctive voice of the

scientific community to those of governments and inter-governmental bodies that have made the case for open data as a fundamental pre-requisite in maintaining the rigour of scientific inquiry and maximising public benefit from the data revolution in both developed and developing countries. It builds on ICSU's 2014 statement (ICSU, 2014) on open access by endorsing the need for an international framework of open data principles.

In the months ahead, partner organisations will promote discussion and adoption of these principles by their respective members and by other representative bodies of science at national and international levels, asking that organisations review the accord and endorse it.

The data challenge for sustainable development

The issue of data – as it is presented in Agenda 2030 (United Nations, 2015) – is identified as playing a fundamental and necessary role in the sustainable development space. The breadth of the 2030 Agenda for Sustainable Development and on-going development of ICTs provide opportunities for a more systematic and ambitious approach to data collection, management and integration.

There is a realisation that many of the indicators, targets, and goals will require monitoring and measuring in order for improvement and eventual achievement. This will draw on a very diverse set of socio-economic and environmental data as recognised in the *ICSU-ISSC Review of the Sustainable Development Goals* (ICSU, 2015).

Central to this realisation is the fact that data will be used to guide decision-making and underpin evidence-informed policy action. Open data improves accountability. Open data is a tool for development.

The need to make decisions based on the evidence and the data means not all countries and state actors will be best equipped to do so, creating a potential divide in progress for the global goals. Less economically developed nations will require investment in infrastructure and capacity for a data-driven process.

Data capacity is something that is unequal across regions. The emergence of new data infrastructures, policies to promote open data and the adaptation of science systems to dealing with data is therefore just as critical (if not more so) in developing regions as it is in high-income countries.

There is great potential for data-driven development because of the detail of data available, the potential timeliness of data availability and analysis, and ability of data to be utilized for multiple purposes at scale. Improved data collection and use of innovative data sources can also make large portions of low-income populations more visible. Although many well-funded national science systems are adapting rapidly to seize the data challenge, the great promise of 'Big Data' (the analysis at scale of ubiquitously gathered unstructured data) remains remote for many less affluent countries, and especially for the least developed countries (LDCs), where the costs of adaptation referred to in the next section pose particular problems.

Unfortunately, the 2030 Agenda foresight does not extend its analysis beyond the needs for improved data collection. Downstream data needs are neglected (storage, analysis, openness etc. are all missing). The national interest and ownership is mentioned but how this pertains to openness of the data is ignored. Public-private cooperation is encouraged, but there is no mentions of the pitfalls of such arrangements (e.g. data used for other purposes and data monopolies).

Given the interlocking nature of the goals and targets, openness of data and capacity to analyse complex and diverse datasets will be key, together with a more comprehensive

analysis of the where the data does not yet exist or is unreliable.

The future of the sustainable development agenda will be data-intensive and its ultimate success will depend on international ability to mobilise the data revolution. This requires working across multiple scales (integrating international, national, local, and city-wide agendas). Considerable responsibility will fall upon researchers operating within the context of local, regional and global multi-stakeholders frameworks to provide the evidence base: but this will not be possible unless the data are openly available and the previously mentioned infrastructure and capacity needs are addressed.

Regional focus on open data platforms that will coordinate and mobilise regional data science capacity to address the dual issues of "Big Data" and "Open Data" will be essential to address the challenges and opportunities for sustainable development.

References

ICSU (2015) Open Data in a Big Data World, An International Accord (http://www.icsu.org/science-international/accord/open-data-in-a-big-data-world-long)

ICSU (2015) Open Data in a Big Data World, An International Accord Abbreviated Version (http://www.icsu.org/science-international/accord/open-data-in-a-big-data-world-short)

ICSU (2014) Open access to scientific data and literature and the assessment of research by metrics (http://www.icsu.org/general-assembly/news/ICSU%20Report%20on%20Open%20Access.pdf)

Royal Society (2012) Science as an open enterprise (https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/)

UNEAG (2014) A World that Counts: Mobilising the data revolution for sustainable development (http://www.undatarevolution.org/report/)

United Nations (2015) Transforming Our World: The 2030 Agenda for Sustainable Development (https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf)

ICSU (2014) Review of Targets for the Sustainable Development Goals: The Science Perspective (http://www.icsu.org/publications/reports-and-reviews/review-of-targets-for-the-sustainable-development-goals-the-science-perspective-2015/SDG-Report.pdf)