Science and the digital revolution: data, standards and interdisciplinary integration

The digital revolution of the past two decades offers profound opportunities for science to discover hitherto unsuspected patterns and relationships in nature and society, on scales from the molecular to the cosmic, and all in areas of human concern, from cultural artefacts and local health systems to global sustainability.

Although there has been considerable progress in a number of fields (e.g. climate science, genomics, astronomy), a major potential remains unrealised: to merge and integrate the data from **different** disciplines in order to reveal deep patterns in the multi-facetted complexity that underlies most of the major global challenges that humanity currently faces. This failing arises from the varying and incompatible standards that have been used across the disciplines of science to codify data, and, in some, inadequate definition of the vocabularies needed to categorise them, with the result that the integration of diverse data can generally only be achieved between closely allied fields. Characterising, understanding, and dealing with the complexity inherent in major global challenges will be integral to the mission of the new International Science Council that will come into being in July 2018.

Realising this potential is of great importance for the science of the 21st century, and was the target of a discussion held at the Royal Society on 13-15 November 2017, organised by the Committee on Data for Science and Technology (CODATA) and supported by the International Council for Science (ICSU) and the International Social Science Council (ISSC). It involved a wide range of natural and social scientists, including data scientists, representatives of international scientific unions and associations, groups working on global challenges and producers of major data services. It was agreed that the participants would work together with the broader research community to:

- develop and apply solutions for interdisciplinary data integration;
- pursue this through data integration for major global challenges that can also act as exemplars of the power of its interdisciplinary potential;
- support, in parallel, the development of capacities to realise the potential of modern data resources across all the disciplines of science; and
- recognise that in many disciplines, foundational work needs to be undertaken to develop the specific vocabularies that are needed to enhance data discovery, use, interoperability and integration.

The participants recognised that these are long-term, decadal, objectives, and that the development of effective momentum and capability to address them would be best served by prioritising work on specific global challenges that would be most tractable in benefitting from better data integration. These would also serve as proofs of concept to motivate further efforts in improving data discovery, access, and reusability in support of interdisciplinary research. The choice of topics built on preparatory work at a meeting held at ICSU HQ in Paris in June 2017. The group will now work together with relevant domain experts to develop proposals for major cross-disciplinary data integration projects to advance solutions for three important global challenges in **infectious disease**, **sustainable cities**, and **disaster risk reduction**.

Addressing these issues requires federated access to distributed data (some key data are required by all three challenges). In each case, the pilot projects will be refined in the coming months, and involve the establishment of project development teams with participants from institutes and organizations, including the communities of ICSU-ISSC international programmes such as *Integrated Research on Disaster Risk*, *Urban Health and Wellbeing*, and *Future Earth*, which are able to provide relevant data and in whose best interest it is to make use of shared, open data.

While these pilot projects will not lead to complete solutions, they will determine the resource requirements and further investment needing to be made to address these and other global challenges, including many that are central to achievement of the UN Sustainable Development Goals. ICSU-CODATA is eager to engage others in the research community who are interested in joining and supporting these challenging projects, the outcome of which could lead to major improvements in the quality of life, human health, and sustainable development worldwide.