Summary Report of the 11th Meeting of the WDS Scientific Committee (WDS-SC)

Jaypee Vasant Continental Hotel, New Delhi
06–07 November 2014

Goals of the Meeting

Topics of focus for the 11th Meeting of the WDS-SC, were:

1. The current role of ICSU-WDS in Less Economically Developed Countries (LEDCs), and the means by which this might be strengthened in the future.

2. Revision of the WDS Data Policy in response to the open access statement endorsed by ICSU at its 31st General Assembly (30 August–04 September; Auckland, New Zealand), and the increase in disciplinary coverage of WDS membership.

3. Examination of the outcomes of the SciDataCon 2014 Conference (02–05 November; New Delhi, India) and the necessary steps for SciDataCon 2016.

ICSU-WDS' Role in LEDCs

- The ICSU World Data System (ICSU-WDS) has completed a project to develop a network data centre for socioeconomic data in Africa. The objectives of the project were: development of a prototype portal to serve as an aggregator for network data contributors, acquisition of metadata from test members, and creation of membership criteria. The project fell short in engaging prospective membership for metadata exchange since the African sources approached were not a in position to supply metadata.

- The project has rewritten the WDS Certification Criteria in such a way that the critical parts are achievable and the remainder relaxed or their responsibility given to someone else. In particular, sustainable operation cannot be established without a trusted digital repository as a core (hub) node that satisfies many of the WDS Criteria and that provides a hardware/technology/regulatory backbone to the satellite nodes of the network. Governance is then a hybrid between resources provided by the hub node and the voluntary members of the network community.

- Data categorization—using a combination of where these are, how they are preserved, and their quality regime—means that only data at a sufficient level are shared within ICSU-WDS—the others will be known but not published—and metadata are captured directly in software provided by the hub node.

- A portal has been generated to submit metadata and point to data files. Automated harvesters can also be implemented for inclusion in a trusted digital repository, such that querying and analyses of the data, and download facilities, are provided. The portal environment can be linked to any number of nodes, and nodal data centres can be very easily and inexpensively cloned as many times as one wishes.

- The next steps are to try to build one of the data centres and populate it. ICSU-WDS will apply for funding to prove the hypothesis, and then approach organizations for philanthropic contributions towards setting up network data centres in LEDCs. However, long-term funding is an issue, and a journal publisher has been approached about the possibility of network data centres being supported as a by-product of an open-access journal model that uses submission fees to safeguard data availability.

- By providing nodal networks in LEDCs, ICSU-WDS will try to address the issue in developing countries of preservation and stewardship of longtail data. The WDS model adds a project landscape around the issue of longtail data, and ICSU-WDS can support its Members to cultivate services capable of dealing with the longtail by demonstrating the important role of domain data services if data are to be discoverable and reusable. Moreover, ICSU-WDS will explore whether it might supply an organizational, community activity for data rescue, limited to the preservation of small datasets in domain repositories.

- If a network data centre satisfying the lightweight criteria reaches a point where it has sustainable funding and it wishes to join ICSU-WDS, then it should apply as a Regular Member since this role is prevalent despite their network function (the contribution of the satellite nodes will be appropriately recognized). Regular Member is not considered as definitely being the correct WDS classification, and the WDS-SC will watch a centre's evolution , and make changes as necessary.

Full & Open Access

- ICSU-WDS will look to create a new set of data-sharing principles that are interoperable/compatible with other initiatives, are adoptable by WDS Members already bound by other policies, and resonate with potential WDS Members. These principles will be broad and highly nuanced since ICSU-WDS has a wider scope of activities than many initiatives, and thus various types of access must be covered. Moreover, exemplars of policies and open licences that support/align with WDS' position will be highlighted.

- A set of draft principles will be drawn up and then fleshed out into a draft policy. These drafts will look to be simple but elegant, clear, reflect WDS' unique value proposition, and align with the history of other bodies. The process will be iterative, and WDS will liaise with the Committee on Data for Science and Technology (CODATA), which has established a Committee on International Data Policy.

- Whilst the charging of fees by WDS Members is exceptional, this might change in the future if data services are threatened with closure because of funding shortages. Moreover, charging fees is a legitimate business model for sustaining data stewardship activities, and must not be mixed up with openness as a principle. Any WDS Member who requests fees should nonetheless make it transparent as to what these are covering such that the WDS-SC can evaluate individual cases.

- 'Cost of Fulfilling User Requests' was defined when data dissemination was physical, and although this marginal cost has dropped towards zero, charging for dissemination has been historically acceptable. However, there are other costs of long-term data stewardship that have increased, and it is not always clear who should pay for them. The discussion is even more complex when curation is involved since it requires a person with high-level expertize to work over a period of time.

- Potentially refining the definition of Cost of Fulfilling User Requests could account for (at least some of) these charges. Notwithstanding, ICSU-WDS does not wish to go against the commonly held view of cost recovery, and its position is that data are openly accessible at minimum charge (for free) as a general rule. There cannot be a situation where users effectively pay twice for tax-funded data.

Communications & Outreach

- With the 11th Meeting following directly after SciDataCon 2014, and in advance of a discussion with the Executive Committee of CODATA, the WDS-SC took the opportunity to review the Conference and to consider the steps needed for SciDataCon 2016.
  - o The SC would like to gain a critical mass of attendance for SciDataCon 2016 Conference, and the location must be carefully selected to achieve maximum participation. The process to host the 2016 Conference will begin in early 2015, and strong bids that bring together Members from both CODATA and WDS will be encouraged.
  - o SciDataCon 2014's broad theme led to some excellent sessions; in particular, the Citizen Science

and Health Data Sessions should be repeated at the next Conference. However, the streams will be thought about more carefully for SciDataCon 2016; for example, there might be a Science/Research Stream, an Applied/Practical Stream, a Task Group Stream , and so on.

- o Linking of different communities was largely successful at SciDataCon 2014, and the 2016 Conference should look to expand on this by involving the science-user community and statisticians, as well as organization such as the Group on Earth Observations (GEO), the United Nations, and those involved with disaster management.
- o SciDataCon 2016 should strive to be the niche Conference that data scientists attend because they find other innovative data scientists there; the natural environment for communities in all scientific domains that use large datasets, and who are not necessarily well-served by their domain meetings.
- o The Conference should also be an exhibition for vendors providing software and hardware components used by the community to (1) showcase the possibilities that come from using their systems, and (2) hear what the community wants to achieve such that they can instigate changes or state their limitations.
- o SciDataCon 2016 needs a greater effort to reach out and involve early-career researchers (ECRs)and PhD students. This might be achieved by adding an age-balance requirement for each session, a specific Plenary Session/side event, or a more interactive Poster/Presentation Session. ECRs may have funding issues to attend the Conference, and a solution for this must be found.

- The WDS-SC also appraised the WDS Members' Forum that took place before SciDataCon 2014.
  - o This event was viewed as a success. The feedback received was generally positive, with several Members stating that interacting with fellow Representatives was especially helpful to understand the community better.
  - o The Plenary Session was useful in informing, and receiving comments from, the community as to the direction ICSU-WDS is heading. Moreover, the talks given by Member Representatives during this Session were very interesting. Future Plenary Sessions might also include time to deliberate a WDS-specific topic of relevance to a large proportion of Members.
  - o The Poster Session did not go as planned since the International Programme Office (IPO) was not aware of the logistical restrictions. This needs to be more carefully checked for the next Forum.
  - o Members expressed that having the WDS logo (and website link) on their homepages is valuable in demonstrating that an external assessment by peers has commended their work. ICSU-WDS will look to reinforce this by showing an elevated statement explaining what it means to be a WDS Member when a user hovers over/clinks on the WDS logo on a Member's website. In addition, the IPO will support WDS Members by providing information that can be used for marketing purposes.
  - o An ad hoc side meeting on the WDS Metadata Catalogue (MC) was deemed highly effective owing to the large attendance of Member Representatives. The response was excellent, with a number of Representatives volunteering their assistance as liaisons to draw from their experiences and to create a more formal Working Group (WG).

- There was consensus by the WDS-SC to endorse the Second International Polar Data Forum proposed by two WDS Members, in light of the importance in keeping polar data at the forefront of people's minds. Especially, the Forum is well integrated with the aims of ICSU-WDS. Nevertheless, the WDS-SC belief was that the Forum's scope must be broadened to more explicitly involve the entire community. A number of practical issue are yet to be decided and must be returned to.

- The WDS-SC unanimously approved a proposal to co-host a SCOSTEP–WDS VarSITI Workshop in September 2015. The WDS-IPO will support this activity, and SC members and WDS Members will be recruited for the Scientific Organizing Committee.

WDS Administrative Matters

- The third WDS-SC will be appointed by the ICSU Executive Board (EB) in spring 2015. There will be

mandatory renewal of five SC members who have served two full terms, with the remaining members put up for re-nomination, if they wish. However, the number of seats on the WDS-SC must return to twelve after the renewal. Since the leadership and members started at the same time, the Chair and Vice-chairs are unfortunately all being lost at once. A simple solution for increasing leadership continuity in the future is to introduce a stagger between the leadership and member rotations. For the current case, the SC unanimously agreed to request the ICSU EB that the Chair be extended as a thirteenth member until after the 13th SC Meeting in order to explicitly handover to the next Chair.

- The current agreement between ICSU and National Institute of Information and Communications Technology (NICT) for the latter to host the WDS-IPO ends in March 2016. Both organizations are satisfied with the performance of the IPO, and their presidents are in the process of negotiating an extension of its hosting within NICT. The SC would like to increase recognition of the IPO by Japan's policy makers and scientific community, and more formal endorsement of the IPO will be sought from other Japanese ministries, and collaborative efforts with the Japanese GEO activity will be explored.

- ICSU-WDS would like to place all of its Member profiles into the re3data registry. The majority of WDS Regular Members are already included—WDS accreditation is one of the recognized certificates—and a recent update of re3data's metadata schema provided ICSU-WDS an opportunity to ensure that all types of data services can be easily included. In 2015, re3data plan to supply a web-interface that will give the IPO an efficient workflow between the WDS membership application process, and registration and management of Member profiles in the re3data registry. Longer term, ICSU-WDS would like to push Member entries into the registry via its Membership Management Tool, and to augment/edit these through a provided application programme interface.

- Under their Charters, WDS-WGs must be renewed every two years. With the renewal deadline of all WGs reached, the WDS-SC made a decision on the continuation of each:
  o The Data Publishing WGs have a timescale under the Research Data Alliance (RDA) of 18 months. The WDS-SC agreed that these WGs will almost certainly continue for the next five years or more.
  o The Certification WG also has an 18-month timescale under RDA, and is set to continue until December 2015. The WDS-SC will thus examine the WG's renewal in more depth at its 13th Meeting.
  o The Knowledge Network (KN) and MC initiatives will be merged into a single KN-WG with the aims of (1) coordinating establishment of WDS infrastructure components in collaboration other stakeholders, and (2) creating a global registry of trusted data services alongside re3data/DataCite.

- The first draft of a Charter for a WG on data-/evidence-based policy making will be written, outlining the advantages and risks, interested projects, and how the work would facilitate collaboration. The Charter will focus on manageable, concrete examples relevant to ICSU-WDS, with the first step to build a conceptual model and identify the required resources to examine the project's feasibility. Future Earth and ICSU's Urban Health and Wellbeing Programme will be included as exemplars, and links with ICSU's Integrated Research on Disaster Risk also studied.

- The WDS-SC examined the possibility of establishing a Task/Discussion Group with CODATA on how to transition from noisy data collected by citizens to data of scientific value that can be stewarded by WDS Members—ensuring a clean data stream from the original raw data to those archived long term. The issues are sociological as well as technical, and this Group would be under CODATA's leadership with WDS contributing technical aspects. The Group might include people from disciplines with maturity in this field, in addition to reaching out to citizen science activities and advocating good practices.

- The 12th Meeting of the WDS-SC will take place on 07–08 March 2015, the weekend immediately before the Fifth RDA Plenary, and will be held at Scripps Institution of Oceanography in San Diego, USA.