CERTIFICATION OF WORLD DATA SYSTEM FACILITIES AND COMPONENTS (23 OCT 2010)

Introduction

A new ICSU World Data System (WDS) has been created by a decision of the 29th General Assembly of the International Council for Science (ICSU). WDS builds on the 50-year legacy of the ICSU World Data Centre system (WDC) and the ICSU Federation of Astronomical and Geophysical data-analysis Services (FAGS). Approximately 100 WDCs and Federation Services, as well as numerous other data centres, services and activities, have already expressed interest in becoming part of the new system. The WDS concept aims at a transition from existing stand-alone WDCs and individual Services to a common globally interoperable distributed data system, incorporating emerging technologies and new scientific data activities. The new system will build on the potential offered by advanced interconnections between data management components for disciplinary and multidisciplinary applications.

The WDS will enjoy a broader disciplinary and geographic base than previous ICSU bodies and will strive to become a worldwide 'community of excellence' for scientific data. To this end, WDS will work closely with ICSU's Committee on Data for Science and Technology (CODATA) and with the new ICSU Strategic Coordinating Committee for Information and Data (SCCID).

Evaluation Criteria

As part of the process of developing the WDS, a transparent, objective base for the evaluating WDS candidates for membership into the WDS data system has been developed. This will ensure the trustworthiness of WDS facilities in terms of authenticity, integrity, confidentiality and availability of data and services. The certification process is based on a catalogue of evaluation criteria (see Table 1). Guidance will be provided for each of the criteria. The procedure to be applied to candidates will be a demonstration of capabilities by means of a questionnaire based on the catalogue of criteria. It can also be used for periodic assessment and monitoring of the WDS facilities and for the overall performance of the WDS. All steps involved are overseen by the WDS Scientific Committee (WDS-SC), which functions as a certificate authority. If necessary, the WDS-SC will engage further experts for the evaluation of specific criteria.

The catalogue of criteria comprises four sections covering policies, organizational framework, management of data, metadata, and services, and technical infrastructure. However, as developments are fast and the scope of WDS comprises a number of special services (former FAGS) as well as new services like data publishing, the catalogue and procedures described may evolve with time. In addition, depending on the type and level of the service provided, specified criteria and procedures may be mandatory, recommended, or optional.

The WDS is structured as a federated system integrating various facilities with different scope. Accordingly, interested parties can be certified and accredited for one or multiple roles. These could include, for example, data collection and processing (including quality assurance), long-term data repository (e.g. data library), data publisher (including periodic compilation of data products), community related service, data analysis service. In other words, in applying the criteria, WDS must take into account the context of the institution, its mission, priorities, and stated commitments. Certification should be seen as an iterative process leading to stepwise improvements and reflecting the overall development of the WDS components. Feedback will be provided to the candidates as part of the process.

The criteria and procedures have been defined incorporating existing standards and best practices from other organizations and projects (OAIS[1], OCLC[2], NESTOR[3], WMO-IS[4], CRL[5], DSA[6]), whilst taking a pragmatic approach balancing proper functionalities against feasibility.

Metrics (measurability)

In principle, the goal is to have objective controls (criteria) against which candidate WDS members can be evaluated. In some cases, questions can simply be answered as yes or no. In many cases, however, evaluation will be based on non-standardized information supplied by applicants indicating the degree of trustworthiness and adequacy. In general, the information supplied for a specific criterion can be attributed to different levels of maturity (for example: not addressed, conception phase, implementation phase, operational). The decision about what is a valid concept or which service can be seen as fully operational is in the responsibility of the WDS-SC, with input from other experts, as appropriate.

A further provision is the obligation level: mandatory (M), recommended (R), or optional (O). This level varies with the different roles within WDS, thus supplying a baseline of minimum requirements per role.

Certification procedure

WDS Certification will be a five stage process:

1. Facility responds to initial WDS survey, or provides a letter of interest
2. Facility demonstrates its capabilities using the questionnaire to describe its capabilities (and possibly by practical demonstrations)
3. If necessary, an on-site review may take place (to be decided by negotiations with the candidate)
4. Accreditation as a WDS component
5. Review of accreditation should take place every 3-5 years

---

[1] Reference Model for an Open Archival InformationSystem (OAIS) - (corresponds to ISO 14721:2003), http://public.ccsds.org/publications/archive/650x0b1.pdf
[2] Trustworthy Repositories. Audit & Certification: Criteria and Checklist, http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf
[3] Dobratz, S et al (2009), Catalogue of Criteria for Trusted Digital Repositories, nestor materials, Deutsche Nationalbibliothek, Frankfurt (Main), Germany, http://nbn-resolving.de/urn:nbn:de:0008-2010030806
[4] http://www.wmo.int/pages/themes/wis/index_en.html
[5] http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/core-re
[6] http://www.datasealofapproval.org/

WDS monitoring

The WDS-SC will be responsible for designing and implementing mechanisms to monitor the overall performance of the system as well as the performance of member facilities.

Table 1: Catalogue of Criteria for WDS Certification

1. WDS general requirements and policies (Organization specific requirements)
1.1   Signed Letter of Agreement, Intent to Cooperate or similar with ICSU
1.2   Have relevant external experts to provide advice and guidance to WDS node
1.3   Should attend WDS bi-annual meetings
1.4   Promote active communication with research community and other users
1.5   Provide full, open, timely, non-discriminatory and unrestricted access to metadata, data, products and services, no cost or at the Cost of Fulfilling User Request (COFUR).
2. Organizational framework
2.1   The facility has defined: (a) the scope of the data and/or product (services) it offers; (b) its responsibility for the long-term preservation its data, products and services; (c) its target user communities and their needs; (d) the rights of its users to access and use data; and (e) processes for responding to changing scientific requirements and to evolving technologies
2.2   The organizational form is adequate for the facility in terms of funding, sufficient numbers of qualified staff, organizational structure and long-term planning
2.3   Expertise of the host organisation offers local oversight (scientists, data specialists) of international repute
2.4   Maintenance of a continuity plan in the event of a host institution shift of interests or reaction to substantial changes
2.5   Facility is committed to formal, periodic review and assessment to ensure responsiveness to scientific and technological developments and evolving requirements
3. Management of data, products and services
3.1   The facility ensures integrity and authenticity of data sets during ingest, archival storage, data quality assessment and analysis, product generation and access and delivery
3. 2   The facility accepts data sets from its producers based on defined criteria for collection, selection and evaluation
3.3   Archival storage of the data sets is undertaken to defined specifications
3.4   The facility permits efficient usage of archived data sets, products and services based on defined criteria and preferably open standards (searchable, accessible, and usable objects and services)
4. Technical infrastructure
4.1   Facility functions on well-supported operating systems and other core infrastructural software
4.2   Facility is using hardware and software technologies appropriate to the services it provides to its designated community(ies)
4.3   Security: Technical infrastructure for protection of the facility and its users, data, products and services