DSA–WDS Partnership: Streamlining the landscape of data repository certification

Rickards, Lesley; Vardigan, Mary; Dillo, Ingrid; Genova, Françoise; L'Hours , Hervé ;
Minster, Jean-Bernard; Edmunds, Rorie; Mokrane, Mustapha

document version
Publisher's PDF, also known as Version of record

Link to publication in KNAW Research Portal

DSA–WDS Partnership: Streamlining the landscape of data repository certification

Authors: Lesley Rickards, Mary Vardigan, Ingrid Dillo, Françoise Genova, Hervé L'Hours, Jean-Bernard Minster, Rorie Edmunds, Mustapha Mokrane

Conference paper

Summary

The Data Seal of Approval (DSA) and the International Council for Science's World Data System (ICSU-WDS) have both developed core certification standards for trustworthy digital repositories and offer their own certification services. However, whilst the DSA and WDS core certifications standards evolved independently, their catalogues of criteria and their review procedures are based on the same principles of openness, transparency, and striking the right balance between the simplicity and robustness of the work and effort involved. A DSA–WDS Partnership Working Group (WG) was initiated under the umbrella of the Research Data Alliance (RDA) with the objectives of realizing efficiencies, simplifying assessment options, stimulating more certifications, and increasing impact on the community. The WG delivered a harmonized DSA–WDS Catalogue of Common Requirements for core certification of repositories, as well as a set of Common Procedures for their assessment. This unified certification standard will be adopted by the two organizations, and integrated into relevant tools and systems, before the end of 2016.

Emergence of standards for certification of data repositories

Data created and used by scientists should be managed, curated, archived, and made accessible in such a way to maximize the initial investment in collecting them by enabling reuse and verifiability of scientific findings. Researchers and other users must be certain that data held in repositories remain useful and meaningful into the future. National and international funders are increasingly likely to mandate open data and data management policies that call for the long-term stewardship and accessibility of data, with some making this a required element in Data Management Plans. Some funders and science publishers go as far as to stipulate that data generated through the research they fund, or which underpin scientific publications, must be deposited in a trustworthy repository.

In response to these requirements, a number of certification standards and accreditation procedures to assess the trustworthiness of data repositories have been developed worldwide in recent years. In addition to the DSA[1] and WDS[2] core certification approaches, the network of expertise in long-term storage and accessibility of digital resources in Germany (nestor) Seal/German Institute for Standardization standard 31644[3] and the Trustworthy Repositories Audit and Certification criteria/International Organization for Standardization (ISO) standard 16363[4] have gone beyond this foundational level in terms of complexity, compliance processes, and investment. For historical reasons, the primary communities served by DSA and ICSU-WDS have been different: the Humanities and Social Sciences by DSA and the Earth and Space Sciences by ICSU-WDS. Nevertheless, both initiatives have fully multidisciplinary missions; they have both been developing towards other communities, and even have common members within their constituencies.

Benefits of core certification

Core certification is designed as a minimally intensive process whereby digital repositories supply evidence that they are trustworthy and have a long-term outlook. A repository first conducts an internal self-assessment, which is then reviewed by community peers. Such assessments help data communities—producers, repositories, and consumers—to improve the quality and transparency of their processes, and to increase awareness of and compliance with established standards. This community approach guarantees an inclusive atmosphere in which the candidate repository and the reviewers closely interact.

In addition to external benefits, such as building stakeholder confidence, enhancing the reputation of the repository, and demonstrating that the repository is following good practices, core certification provides a number of internal benefits to a repository. Specifically, core certification offers a benchmark for comparison—especially, when evidence statements are publically available—and helps to determine the strengths and weaknesses of a repository.

The applicants to DSA and ICSU-WDS generally see such a certification as beneficial, since not only does certification support trust in the repository, it also stimulates a repository to improve its processes and procedures and move to a higher level of professionalism, with an incentive to enhance its operations over time. Finally, core certification offers a solid foundation for the repository to apply for a higher-level certification in the future.

Harmonizing the DSA and WDS core certifications

In 2013, the Repository Audit and Certification WG[5]—including representatives from the DSA and WDS communities—was established under the umbrella of the RDA–WDS Interest Group on Certification of Digital Repositories to nurture a DSA–WDS Partnership, with the objectives of realizing efficiencies, simplifying assessment options, stimulating more certifications, and increasing impact on the community.

On the basis of a side-by-side analysis of their certification frameworks, the two organizations have moved forwards over time to unified wording and criteria. The WG delivered a harmonized Catalogue of Common Requirements[6] for core certification of repositories that is drawn from the DSA and WDS criteria, as well as a set of Common Procedures[7] for their assessment. The Requirements include an introduction on the benefits of certification, background information, guidance text, and a glossary. Both the new Catalogue and the Procedures—to some extent—have been validated through a testbed comprised of DSA and WDS community members undertaking test assessments. The testbed results have been published and the insights resulting from this and subsequent minor improvements to the Common Requirements will be assessed.

This presentation will examine the Common Requirements; outlining the issues overcome in establishing the Catalogue, and discussing the auditing process and—in particular—how to reconcile self-assessment and objective measurement of compliance with the criteria.

Future of core certification

By uniting the fundamental and lightweight certification mechanisms of DSA and ICSU-WDS, the joint WG has taken a first step in simplifying the array of certification options, and has shown the value to be gained from a certification procedure requiring relatively low investment of time and effort.

The DSA–WDS Common Requirements and Procedures are expected to replace the current certification catalogues of both organizations before the end of 2016. It is expected that the new DSA–WDS certification standard will get a boost in adoption by creating a critical mass of certified repositories. Becoming certified through either organization means the repository is certified in the other, eliminating duplication of effort. Already certified repositories will transition to the new DSA–WDS Common Requirements as they renew their certifications. Finally, the biggest potential impact of this work is that it is a first step towards having a framework for more coherent, increasingly stringent, compatible and comparable standards for repository certification, as shown in **Figure 1**.
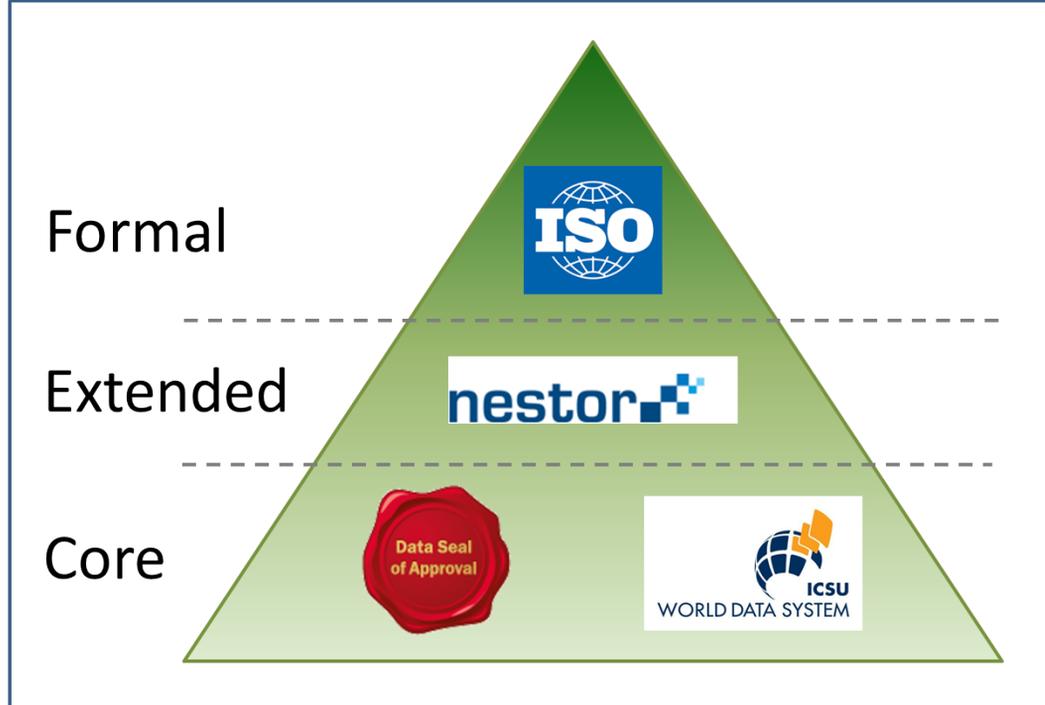
Figure 1: The different certification levels, each with increasing rigour. DSA and ICSU-WDS represent core certification.

Looking to the future, a funding proposal—Trust4Data—that builds on the WG outputs has been submitted to the European Union. Its objective is to establish a Core Certification Service for a wide range of research data repositories. It carries the ambition to develop and create a standard information (object and process) model for core certification and to deploy this in a service within a dedicated group of early adopters. A brief overview of the proposal and outputs will be given.

Acknowledgements

Competing Interests

The authors declare that they have no competing interests.

Notes

1 www.datasealofapproval.org/en/ (http://www.datasealofapproval.org/en/)

2 www.icsu-wds.org/services/certification (http://www.icsu-wds.org/services/certification)

3 http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html (http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html)

4 https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/iso16363 (https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/iso16363)

5 https://goo.gl/TgDvoR (https://goo.gl/TgDvoR)

6 https://goo.gl/jx2ONC (https://goo.gl/jx2ONC)

7 https://goo.gl/Q2xDOp (https://goo.gl/Q2xDOp)