Geophysical Research Abstracts
Vol. 21, EGU2019-19049, 2019
EGU General Assembly 2019

Building Trust in Scientific Data: CoreTrustSeal Certification and the World Data System

Rorie Edmunds (1)

(1) World Data System International Programme Office, Tokyo, Japan, (2) CoreTrustSeal Standards and Certification Board

Data repositories are increasing valued as a key element of global research infrastructure, playing a central role in the long-term preservation of research data that continue to escalate in volume and diversity. Scientific integrity and norms dictate that data created and used by researchers should be managed, curated, and archived in a data repository to ensure that science is verifiable and reproducible while preserving initial investment in data collection. However, to guarantee generated datasets remain available, useful, and meaningful into the future, research stakeholders—scientists, funders, librarians, and publishers—must be able to establish the trustworthiness of a data repository. The need for trustworthiness, and especially for Trustworthy Data Repositories (TDRs), is therefore recognized as a prerequisite for efficient scientific research and data sharing.

Data repository certification is the process whereby data repositories supply evidence to, and are assessed by, an independent authority for their trustworthiness and sustainability against defined criteria through a transparent and objective procedure. Certification helps data communities—producers, repositories, and consumers—to improve the quality and clarity of their practices, and to increase awareness of, and compliance with, established standards.

Nowadays, certification standards for data repositories are available at three different levels: Core, Extended, and Formal. The Data Seal of Approval and the World Data System historically offered separate core certification standards. Drawing from their respective criteria, and within the framework of the Research Data Alliance, the two communities created and adopted a set of harmonized Core TDRs Requirements: 16 universal guidelines intended to reflect the characteristics of a TDR for certification of data repositories at the core level. The certification standard is administered under the authority of a new entity ‘CoreTrustSeal’, whereby a Standards and Certification Board of community representatives grants certification after peer-review of applications based on the Requirements.

Even at this core level, CoreTrustSeal certification offers many benefits to a repository and its stakeholders. It is a minimally intensive process that accounts for the specific aims and context of a repository, and gives it independent insights as to how it may mature and further increase its trustworthiness. Moreover, CoreTrustSeal certification offers a solid foundation if a repository hopes to attain higher-level certification in the future. Self-assessment against the Requirements is useful whether or not a repository wishes to apply for certification, since it enables the repository to appraise its internal procedures with respect to the criteria and to update them where necessary.

Networks (umbrella bodies), which vary greatly in their makeup and remits, are outside the scope of the CoreTrustSeal and their accreditation remains a WDS-only focus. WDS has developed criteria and a procedure to ensure the trustworthiness of its Network Members. Notwithstanding, to enhance the verifiable level of trust in the scientific process and the feasibility of reproduction, many other elements of research activity also require some form of trustworthy service or infrastructure component, and WDS is now exploring the development and provision of core-level certifications to ensure the trustworthiness of such services and components.