Trustworthy Digital Preservation Repositories: an Introduction

Today, the vast majority of information is created and maintained in digital form. Digital preservation is the set of processes by which this information can be made accessible over time. Digital preservation is not just a technological challenge; stewards of digital materials need not just technical systems that will preserve the materials over time, but also organizational and financial frameworks that will ensure their persistence. Challenges include not only storage media that fails or software that becomes obsolete; organizations are susceptible to economic downturns and employee turnover. Safeguards need to be in place to ensure that even in worst case scenarios, digital preservation activities will persist.

How can an institution be assured that it has established the right infrastructure and ecosystems for its digital preservation practices? Several frameworks exist for evaluating and auditing digital preservation repositories and their environments. By using these frameworks as evaluative tools, repositories can assess their levels of compliance and either claim status as a "trusted" digital repository or work to address any uncovered shortcomings. This status is desirable for organizations who engage in long-term digital preservation because it signifies to relevant stakeholders that the repository is a responsible steward of the materials and can be trusted to manage them appropriately. In some contexts, "trustworthy" status has financial impacts as well as reputational, as resource allocators, funding agencies, and data depositors may prefer engaging with digital repositories certified as trustworthy over those without such status.

Trusted Digital Repositories

What does "trustworthy" or "trusted" mean? For a digital preservation system to be considered trustworthy, it must demonstrate that it operates in the manner it specifies according to its objectives and principles. The methods by which systems ingest, preserve, and provide access to digital materials over time are typically reviewed in reference to the Reference Model for an Open Archival Information System (OAIS) standard, or ISO 14721:2012, which provides the language and framework for describing long-term preservation systems and delineates the roles and responsibilities of system participants. The evaluative frameworks discussed later in this article detail the criteria for digital object management practices using OAIS terminology.

OAIS provides the language with which to describe management of digital materials; however, the appraisal of trustworthiness must include an evaluation of the entire ecosystem in which the repository resides. A review of the organization managing the repository is a key piece of this evaluation. How is it governed? Is it appropriately staffed? Are there adequate financial resources allocated to the repository? There must be a demonstrated institutional commitment to preserving a specified set of digital materials over time. This commitment is best exhibited by institutional policies that clearly outline the repository's mission and underscore the importance of preservation to the organization as well as its dedication to preservation practices. Financial statements and dedicated budget lines also can indicate the organization is not only voicing support for its digital preservation systems, but is also dedicating current and future resources to them.

A trusted digital repository must also understand the threats and risks associated with long-term digital preservation and coordinate activities to protect the repository from potential data loss. Digital preservation repositories must be kept secure from external and internal attacks and be monitored for hardware and software obsolescence. Ideally, digital preservation systems replicate data to different geographic locations to protect against natural or manmade disasters. Economic failure is also a threat to preservation systems, and having a succession plan in place may be required for certification. Management of a digital repository is an ongoing series of activities designed to decrease risk; certification requires a review of these activities to ensure the correct actions and functions are present in the system to mitigate against potential loss.

European Framework for Audit and Certification of Digital Repositories

One key aspect of being a trustworthy digital repository is the notion that the criteria for trustworthiness are defined by an outside organization. The criteria also need to be applied objectively. Any organization can declare itself trustworthy, but the designation is meaningless unless the criteria by which it is deemed trustworthy have been clarified. The main frameworks that have been developed to determine trustworthiness are TRAC/ISO 16363, DIN 31644, and the Data Seal of Approval (DSA)[1]. Given that there are several evaluative methods available for repositories to choose from, the European Framework for Audit and Certification of Digital Repositories is an attempt to clarify what "certification" as a trusted digital repository actually means[2]. Representatives from each of the organizations responsible for developing and maintaining the three main evaluative frameworks (TRAC/ISO 16363, DIN 31644, and the DSA) signed an Memorandum of Understanding specifying three levels of auditing certification: basic, extended, and formal. *Basic certification* is granted to repositories that obtain DSA certification. *Extended certification* involves a self-audit based on either ISO 16363 or DIN 31644. This audit must be publicly available and externally reviewed. And, finally, *formal certification* is granted to those repositories that receive both a basic certification and an official external audit on either ISO 16363 or DIN 31644.

---

[1] Another well known evaluative tool is DRAMBORA, though it is used less for determining trustworthiness and more for evaluating risks to digital preservation systems. See http://www.repositoryaudit.eu/

[2] See: http://www.trusteddigitalrepository.eu

Data Seal of Approval

The Data Seal of Approval was originally developed in 2008 as part of the Dutch DANS (Data Archiving and Networked Services) mandate. The seal was recognized as a useful evaluative tool for other data repositories, and in 2009 management and further development of the DSA was transferred to an international body, the Data Seal of Approval (DSA) Board[3].

The DSA consists of 16 guidelines, designed as a minimum set of qualifications distilled from Nestor, DRAMBORA, and TRAC, as well as two additional publications: *Foundations of Modern Language Resource Archives* by the Max Planck Institute and *Stewardship of Digital Research Data: A Framework of Principles and Guidelines* by the Research Information Network. The first four guidelines focus on the data producer, or the stakeholder responsible for the quality of the digital data. Guidelines 4 to 13 outline the responsibilities of the data repository, focusing on the quality of the organizational framework in which the repository resides as well as its technical infrastructure. The last three guidelines concern the data consumer who is responsible for respecting any access regulations, codes of conduct, or licenses applicable to the use of the data. To receive the DSA certification as a Trusted Digital Repository, a repository must show it is compliant with the guidelines pertaining to data repository responsibilities and that it enables data producers and consumers to comply with their respective guidelines[4].

The process for obtaining approval involves submitting an application and self-assessment to the DSA Board. The DSA Board then appoints an outside reviewer to evaluate the application. The reviewer can either approve the application and a Data Seal of Approval is granted, or request additions to the application. Once granted, the Seal can be displayed indefinitely. However, repositories will need to renew their applications to stay compliant with the current standard and receive the latest logo.

DIN 31644

In December 2004 the German nestor (Network of Expertise in Long-term STOrage of Digital Resources) project instituted a working group on Trusted Digital Repository Certification. Consisting of representatives from various German and Austrian cultural heritage organizations, the group published a *Catalog of Criteria for Trusted Digital Repositories* in 2006[5]. A second version of the Catalog was published in 2009. The Catalog eventually was expanded for application to all institutions that preserve digital resources and developed as a German standard (DIN 31644). In contrast to the Data Seal of Approval, which only lists 16 guidelines for evaluation, DIN 31644 specifies 34 metrics for Trusted Digital Repository status. These metrics are divided into three categories: organizational framework, object management, and infrastructure and security. A final version of DIN 31644 was published in German in 2012, and plans for translation are ongoing[6].

ISO 16363

The most extensive of the evaluative frameworks for trusted digital repositories is ISO 16363. In contrast to the 16 guidelines listed in the DSA and the 34 criteria listed by DIN 31644, ISO 16363 details over 100 metrics for evaluation. The development of ISO 16363 is rooted in one of the earliest frameworks by which to evaluate the trustworthiness of digital repositories, the *Trusted Digital Repositories: Attributes and Responsibilities* report published in 2002 by the Research Libraries Group (RLG) and Online Computer Library Center (OCLC). One of the recommendations of this report was to form a process by which repositories could be certified as trustworthy[7]. Following this recommendation, in 2003 the U.S. National Archives and Records Administration and RLG (by then a part of OCLC) formed a task force to address the process of certifying digital repositories as trustworthy. This task force incorporated valuable contributions from the Center for Research Libraries, as well as from the international entities nestor and the Digital Curation Centre, and published the *Trustworthy Repositories Audit & Certification: Criteria and Checklist* in 2007. The Checklist outlines tools for auditing, assessment, and certification of trusted digital repositories. Notably, it also provides examples of the types of documentation a repository can submit to provide evidence of compliance with the requirements listed in TRAC. In 2012, the TRAC Checklist was formalized into the international standard, ISO 16363, Audit and Certification of Trustworthy Digital Repositories.

Outline of the Standard

ISO 16363 is divided into three sections: Organizational Infrastructure, Digital Object Management, and Infrastructure and Risk Management.

The Organizational Infrastructure section covers criteria concerning the organization that manages the digital repository: its governance and organizational viability, organizational structure and staffing, procedural accountability and preservation policy framework, financial sustainability, and, lastly, the contracts, licenses and liabilities surrounding the materials placed in the digital repository.

Digital Object Management concerns the actual handling of the data selected for long-term preservation, including their acquisition, ingest into the system, the creation and management of the archival ingest package, preservation planning, the management of descriptive and identifying information, and the management of access to the system and data objects.

The Infrastructure and Security Risk Management metrics detail requirements a repository needs to manage the risks associated with digital preservation: hardware and software management and monitoring, data fixity monitoring, defined processes for procedures such as storage migrations, and methods for identifying and addressing defined security risks to the system.

Certification Process

To become certified as a trusted digital repository according to ISO 16363, the organization interested in receiving an audit must first contract with an certified auditing agency that has been accredited to conduct ISO 16363 audits by an Accreditation Body, such

---

[3] See: http://datasealofapproval.org/media/filer_public/2014/10/03/20141003_dsa_overview_defweb.pdf
[4] ibid
[5] See: Dobratz, S., Schoger, A., & Strathmann, S. (2007). The nestor Catalogue of Criteria for Trusted Digital Repository Evaluation and Certification. *Journal Of Digital Information, 8*(2). Retrieved from*https://journals.tdl.org/jodi/index.php/jodi/article/view/199/180*
[6] See: http://data-archive.ac.uk/curate/trusted-digital-repositories/standards-of-trust?index=3
[7] See: Research Libraries Group. (2002). Trusted Digital Repositories: Attributes and Responsibilities.Retreived from *http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf*

as the ANSI-ASQ National Accreditation Board in the United States.These auditing agencies must adhere to metrics outlined in the two ISO standards relating to the audit process: ISO 16919, Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories, and ISO 17021, Requirements for bodies providing audit and certification of management systems. The repository and certified auditing agency develop an audit plan, the first step of which is usually a self-audit by the digital repository. The auditing agency reviews the self-audit and then follows up with an on-site visit. The digital repository must produce evidence of conformance to each metric listed in the standard. This evidence can take many forms, including policies, documented procedures, log files, and other documentation. ISO 16363 lists illustrative examples with each metric detailing what types of evidence can be applied to show conformance. Major and minor nonconformances to the standard are reported, and the organization has the opportunity to remedy any issues and submit additional documentation. Barring any major nonconformances requiring additional visits from the auditor, the auditing process leads to certification in about a year. It's important to note that it's not sufficient to undergo the certification process once; there must be a regular schedule of self-assessment and external audits in place. Formal audit renewals should be expected to occur every three years.

Although ISO 16363 was finalized as a standard and published by the International Standards Organization in 2012, there are currently no repositories certified by its metrics. ISO 16919 was published in late 2014 and as of the writing of this article, no Auditing Bodies have been accredited to perform ISO 16363 audits.

Conclusion

Depending on the type of certification desired – basic, extended, or formal – the process of certification as a trusted digital repository can take anywhere from a couple of days to more than a year. Many repositories start with basic certification and then perform a self-audit using ISO 16363 metrics with the intention of undergoing the formal audit process when it becomes available. Whichever path a repository chooses, the process of self- and external evaluation highlights any deficiencies in procedures and processes and communicates to resource allocators and stakeholders the repository's digital preservation capabilities.

Sibyl Schaefer

sschaefer@ucsd.edu ■

(The UC San Diego Library)