Repository Assessment Methods

Umar Qasim

Digital Preservation Officer

University of Alberta

Respository Assessment Helps to

- Promote trust in the long-term care to data
- Improve transparency of the repository
- Improve processes and procedures
- Align with a community standard

Common Assessment Areas

- Organization
  - Governance, staffing, policies, finances, etc.
- Technical Infrastructure
  - System design, security, etc.
- Object Management
  - Access, integrity, process, preservation, etc.

Assessment Options

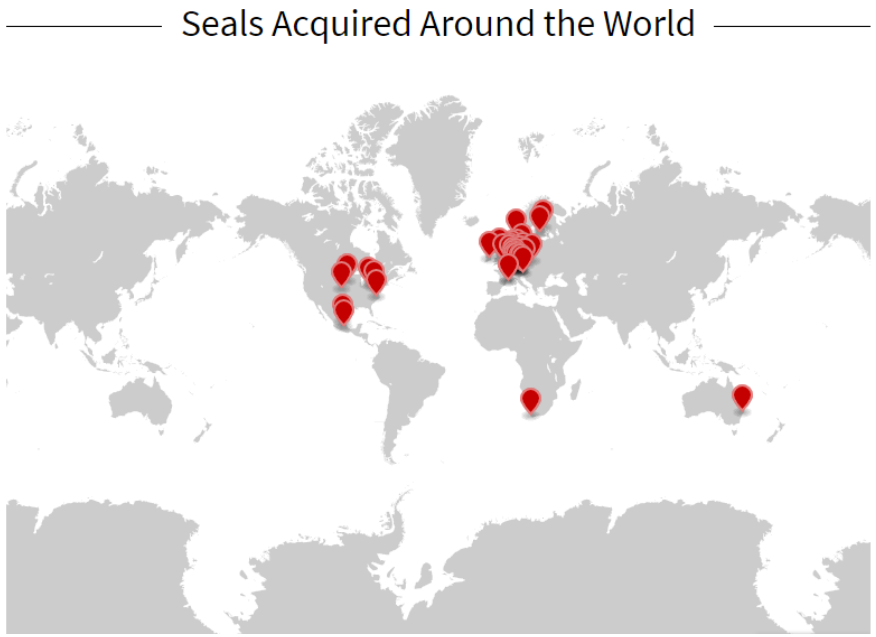
- **Basic Certification**
  - Data Seal of Approval (DSA)
  - World Data System (WDS)
- **Formal Certification**
  - Trustworthy Repositories Audit and Certification (ISO 16363)
  - NESTOR Seal (DIN 31644)
- **Other alternatives**
  - Self-audits (DRAMBORA, ISO 16363, DIN 31644)
  - Peer Review

European Framework for Audit and Certification

- In an effort to coordinate approaches to audit and certification of digital repositories, in 2010 a memorandum was signed to create a “European Framework for Audit and Certification of Digital Repositories”
- The framework integrates three standards: the Data Seal of Approval (DSA), DIN 31644, and ISO 16363.
- A tiered approach by defining three levels of certification suitable to an organization’s size, objectives, and available resources.
  - **Basic Certification:** Obtained through the DSA, a set of 16 guidelines relating to data producers, repositories, and users. To get the DSA, repositories perform a self-assessment using the guidelines.
  - **Extended Certification:** Repositories which have received the DSA can obtain an extended certification by getting an externally reviewed self-assessment based on either ISO 16363 or DIN 31644.
  - **Formal Certification**: Requires that repositories obtain the DSA and get a full external audit done in accordance with either ISO 16363 or DIN 31644.

Source: http://www.trusteddigitalrepository.eu

Data Seal of Approval (DSA)



- Started by DANS in 2009
- 16 guidelines –
  - 3 target the data producer,
  - 3 the data consumer, and
  - 10 the repository
- Self-assessments are done online with ratings and then peer-reviewed by a DSA Board member
- 48 certified repositories

Source: http://datasealofapproval.org/en/

DSA Guidelines

Guidelines Relating to Data Producers:

- The data producer deposits the data in a data repository with sufficient information for others to assess the quality of the data and compliance with disciplinary and ethical norms.
- The data producer provides the data in formats recommended by the data repository.
- The data producer provides the data together with the metadata requested by the data repository.

Guidelines Related to Data Consumers:

- The data consumer complies with access regulations set by the data repository.
- The data consumer conforms to and agrees with any codes of conduct that are generally accepted in the relevant sector for the exchange and proper use of knowledge and information.
- The data consumer respects the applicable licences of the data repository regarding the use of the data.

DSA Guidelines

Guidelines Related to Repositories:

- The data repository has an explicit mission in the area of digital archiving and promulgates it.
- The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.
- The data repository applies documented processes and procedures for managing data storage.
- The data repository has a plan for long-term preservation of its digital assets.
- Archiving takes place according to explicit work flows across the data life cycle.
- The data repository assumes responsibility from the data producers for access and availability of the digital objects.
- The data repository enables the users to discover and use the data and refer to them in a persistent way.
- The data repository ensures the integrity of the digital objects and the metadata.
- The data repository ensures the authenticity of the digital objects and the metadata.
- The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.

Source: http://datasealofapproval.org/en/

World Data Systems (WDS) Certification

- WDS is an effort of the International Council of Science (ICSU)
- Started in natural sciences and similar to Data Seal of Approval
- 20+ criteria (guidelines)
- Membership levels and certification mechanisms(Network, Regular)
- Goal is to enable universal and equitable access to quality-assured scientific data, data services, products and information and to ensure long-term data stewardship
- Foster compliance to agreed-upon data standards and conventions
- Provide mechanisms to facilitate and improve access to data and data products
- Organization completes an Expression of Interest and demonstrates its capabilities using the application form.

Source: https://www.icsu-wds.org

WDS Approach

- The certification processes are based on catalogues of evaluation criteria: one catalogue for Regular Members and another one for Network Members.
- Focuses on—**policies, organizational framework, network framework, management of data, metadata, and services, and technical infrastructure.**
- Incorporates existing standards and best practices from other organizations and projects **(OAIS, OCLC, NESTOR, WMO-IS, CRL, DSA).**

Source: https://www.icsu-wds.org

WDS Certification Criteria

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

WDS Certification Criteria

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

Source: https://www.icsu-wds.org/files/wds-certification-summary-11-june-2012.pdf

the Network of Expertise in Long-Term Storage of Digital Resources (NESTOR SEAL)

- Aimed at German memory organisations and institutions.
- Provides guidance, tools for self-checking, and potentially certification
- *Abstract* criteria, applicable for a range of digital repositories, and valid over a longer period, focus on conformity with OAIS
- Based on DIN 31644 Standard (Criteria for Trustworthy Digital Archives.)
- 34 requirements structured in 3 parts
  - Organization
  - Management of intellectual entities and their representations
  - Infrastructure and security

NESTOR SEAL (DIN31644)

1. The institution wishing to obtain the nestor seal notifies nestor and nominates 2 contact persons
2. NESTOR confirms the start of the review and appoints a reviewer
3. The archives conducts the self assessment and submits it to the nestor reviewer
4. The reviewer checks the plausibility of the self assessment and any accompanying documentation and writes a review report
5. A second reviewer approves the review report
6. Cases of dispute are taken to the nestor Working Group on Certiciation

- The entire process should not take longer than 3 months!

Source: http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html

Trusted Repositories Audit & Certification TRAC / ISO 16363

- Establishes appropriate methodologies for determining the soundness and sustainability of digital repositories
- Provides tools for the audit, assessment, and potential certification of digital repositories
- Establishes documentation requirements required for audit
- A process for certification based on ISO 16363 standard

Source: http://www.crl.edu/archiving-preservation/digital-archives/metrics

TRAC Criteria

Organizational Infrastructure

- 3.1) Governance & Organizational Viability
- 3.2) Organizational Structure & Staffing
- 3.3) Procedural Accountability & Preservation Policy Framework
- 3.4) Financial Sustainability
- 3.5) Contracts, Licenses, & Liabilities

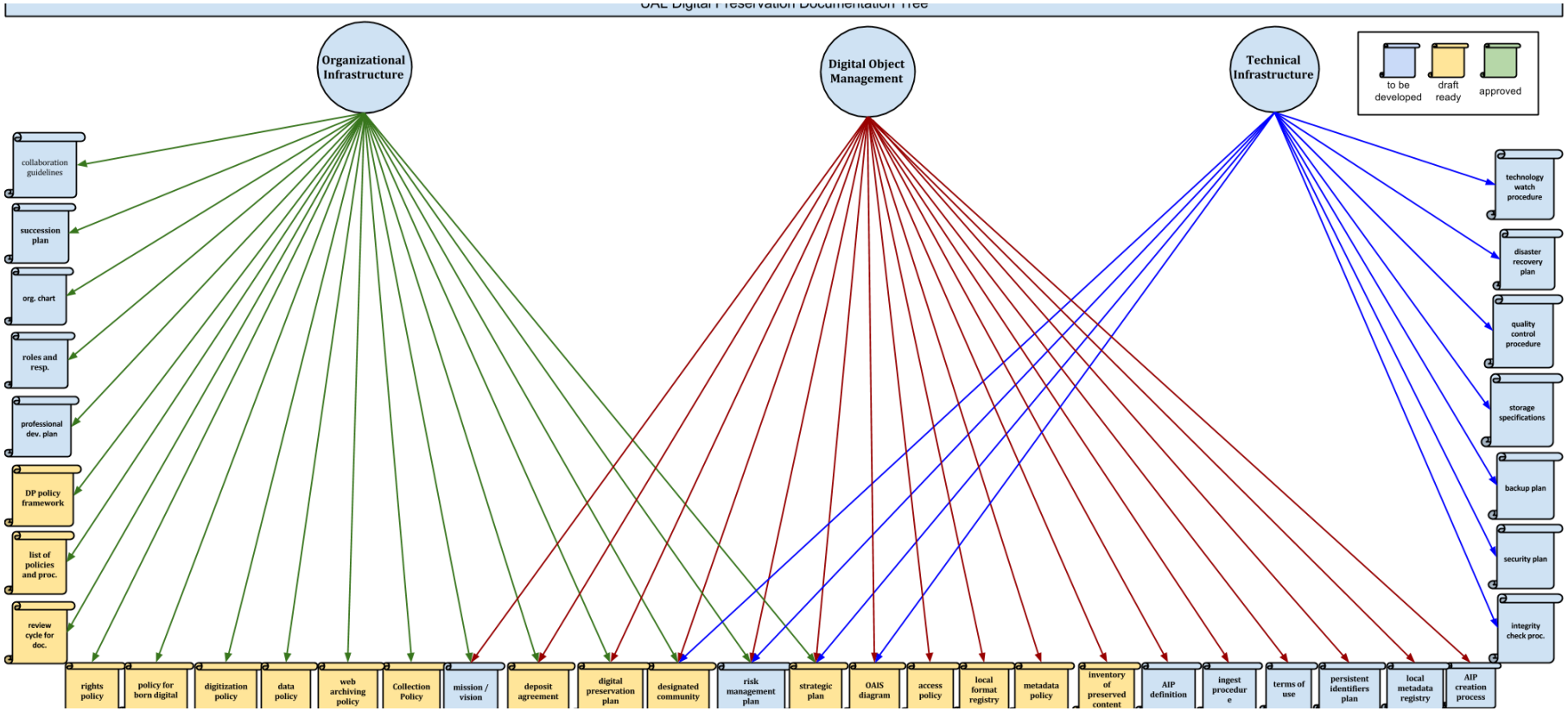
Digital Object Management

- 4.1) Ingest - Acquisition of Content
- 4.2) Ingest - Creation of the AIP
- 4.3) Preservation Planning
- 4.4) AIP Preservation
- 4.5) Information Management
- 4.6) Access Management

Infrastructure and Security Risk Management

- 5.1) Technical Infrastructure Risk Management
- 5.2) Security Risk Management

Sample Documentation Structure



Digital Repository Audit Method Based On Risk Assessment – (DRAMBORA)

- DRAMBORA (Digital Repository Audit Method Based on Risk Assessment) is a self-audit toolkit developed by the Digital Curation Centre (DCC) and DigitalPreservationEurope (DPE) to help guide repository managers to examine and analyse the work of the repository.
- Documented understanding of the risks expressed in terms of their likelihood and potential impact.
- Means for risk management, determining the appropriate strategies for avoidance, treatment, transfer and tolerance, as well as the mechanics of their implementation
- Self assessment with DRAMBORA could be a precursor to external audit, accreditation, and certification when these services become broadly available.
- The six stages of self-audit correspond closely to the preparatory work that organisations will be expected to undertake prior to exposure to full audit.

Source: http://www.repositoryaudit.eu/

Peer Review Services

DPC Peer Review Services

- Provide peer review and reporting on digital preservation facilities and services being offered by members.
- Provide formal evaluation and certification for digital repositories being operated by DPC members.
- Provide means for staff exchange between members to support mutual improvement
- Provide basic quality improvement planning for digital repositories being operated by DPC members

COPPUL Peer Review Services

- Discuss and rationalize local approaches, challenges and issues in developing digital preservation policies
- Review diverse approaches taken by peers in developing digital preservation policies
- Discover and contribute selected resources in the digital preservation community and assess these for relevance to local requirements.

Source: http://www.dpconline.org/about/working-groups-and-task-forces/728-peereview

Questions