

ISO 16363 Self-Assessment at the British Library

Caylin Smith
Repository Manager
caylin.smith@bl.uk
 @caylinssmith

Outline

- ❑ Digital Preservation at the British Library
- ❑ The Library's Digital Collections
- ❑ Achieving trust
- ❑ ISO 16363
- ❑ The Library's ISO 16363 self-assessment

Digital preservation at the British Library



- ❑ Team comprised of 7 people
- ❑ Works cross site between London and Boston Spa
- ❑ Based in Collection Management division
- ❑ Examples of current projects
 - ISO 16363
 - Flashback
 - Preservation of Ingested Collections: Assessments, Sampling, & Action plans (PICASA) project

Our digital collections

Our digital collections

- Born-digital content is predominantly acquired under legal deposit legislation
- Born-digital and digitised content is also deposited voluntarily
- Collection currently comprised of over 12 million items (~550 TB)
- Digital items preserved within the Library's Digital Library System repository and replicated to 4 geographically separate nodes
- Ebooks, ejournals, web archives, newspapers, theses, datasets, audiovisual, maps, electoral registers, personal digital archives, manuscripts, intellectual property documentation, sheet music, stamps, prints, drawings, and photographs

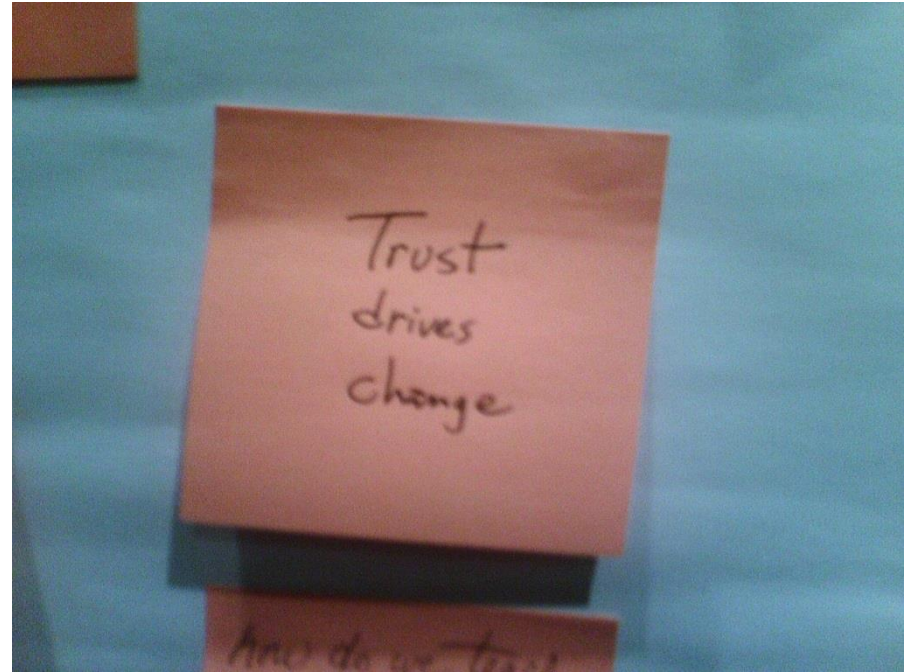
“The British Library is increasingly a digital library.”

Achieving trust

Achieving Trust

“Certification for digital repositories will involve far more than the documentation of criteria...It must recognize standards and best practices relevant to the community of the repository, as well as those of the information management and security industries as a whole. In other words, audit and certification of trust digital repositories cannot exist in a vacuum.” –CRL & OCLC, 2007

“Claims of trustworthiness are easy to make but are thus far difficult to justify or objectively prove. Establishing more clear criteria detailing what a trustworthy repository is and is not has become vital.” ISO 16363, 2011



What factors influence trust?

Stakeholder trust in the organization

- Benevolence
- Integrity
- Identification
- Transparency

Social factors

- Peers
- Mentors
- Senior colleagues

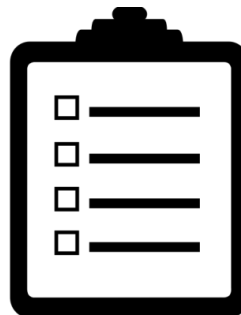
Structural Assurance

- Guarantees
- Reputation
- Third-party endorsement

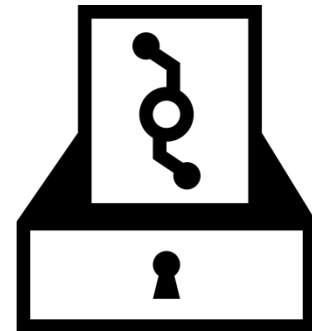
ISO 16363

About ISO 16363

- ❑ Replaced Trusted Repository Audit Certification (TRAC; 2007) in 2012
- ❑ Recognised standard by the digital preservation community
- ❑ References ISO 14721: *Reference Model for an Open Archival Information System (OAIS)*
- ❑ Comprised of 109 metrics divided into three areas: Organisational infrastructure (25), Digital Object Management (60), and Infrastructure and Security Risk Management (24)
- ❑ Defines 'repository' as an organisation responsible for digital preservation, not just the technical system



Created by Aaron Dodson
from Noun Project



The Benefits

□ Self-assessment:

– Gap Analysis

- How are we doing?
- Where must we improve?

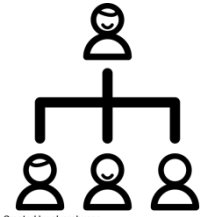
□ Certification:

– Funding Assurance

- We are worth the money (to funding body)
- You can trust us to look after your content (to service users)

Organisational Infrastructure

- ❑ Repository mission statement
- ❑ Repository Strategic Plan
- ❑ Collection Policy
- ❑ Staff who has knowledge to undertake repository activity
- ❑ Succession plan, contingency plans, and/or escrow arrangements
- ❑ Defining the repository's designated community and knowing this community's needs



3.3.1 The repository shall have defined its Designated Community and associated knowledge base(s) and shall have these definitions appropriately accessible.

Supporting Text

This is necessary in order that it is possible to test that the repository meets the needs of its Designated Community.

Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement

A written definition of the Designated Community.

Discussion

The Designated Community is defined as ‘an identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the archive and this definition may change/evolve over time’ (OAIS Glossary, reference [1]).

Examples of Designated Community definitions include:

- General English-reading public educated to high school and above, with access to a Web Browser (HTML 4.0 capable).
- For Geographic Information System (GIS) data: GIS researchers—undergraduates and above—having an understanding of the concepts of Geographic data and having access to current (2005, USA) GIS tools/computer software, e.g., ArcInfo (2005).
- Astronomer (undergraduate and above) with access to Flexible Image Transport System (FITS) software such as FITSIO, familiar with astronomical spectrographic instruments.
- Student of Middle English with an understanding of Text Encoding Initiative (TEI) encoding and access to an XML rendering environment.

Variant 1: Cannot understand TEI;

Variant 2: Cannot understand TEI and no access to XML rendering environment;

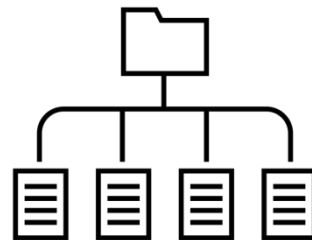
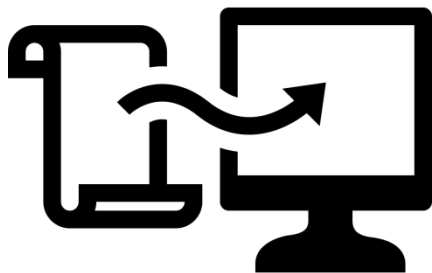
Variant 3: No understanding of Middle English but does understand TEI and XML.

- The repository has defined the external parties, and its assets, owners, and uses. Two groups: the publishers of scholarly journals and their readers, each of whom have different rights to access material and different services offered to them.

Some repositories may call themselves, for example, a ‘dark archive’, an archive that has a policy not to allow consumers to get access to its contents for a certain period of time, but they would nevertheless need a Designated Community.

Digital Object Management

- Deposit and submission agreements
- Deciding packaging information for the SIPs
- Representation Information for the SIP Content Data
- Process that verifies SIPs and AIPs
- Documentation clearly linking each AIP, or class of AIPs, to its definition
- Documentation of the SIP-AIP relationship
- Preservation policy, workflows, and planning
- Written documentation of decisions and/or action taken
- Integrity checking and recovery/self-healing
- Comply with access policy



Created by Juan Pablo Bravo
from Noun Project



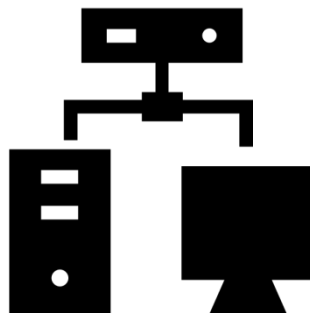
Created by Vicons Design
from Noun Project

Infrastructure and Security Risk Management

- Identify and manage risks associated with the repository
- Technology monitoring and mitigating obsolescence
- Processes for storage media and/or hardware change; upgrades
- Recording and reporting to its administration all incidents of data corruption or loss, and steps shall be taken to repair/replace corrupt or lost data
- Adequate hardware and software support for backup functionality
- Disaster preparedness and recovery plan(s)
- At least 1 offsite backups of all preserved content



Created by Creative Stall
from Noun Project



Created by Creative Stall
from Noun Project



Created by Creative Stall
from Noun Project

The Library's ISO 16363 self-assessment

Objectives

1. Conduct an internal assessment of the Library's approach to preserving digital collections against the ISO 16363 standard to understand how we're doing
2. Identify areas to improve trustworthiness, review with stakeholders, and agree responsibilities for an outline plan of action to address these

Approach Stage 1

- ❑ Become familiar with ISO 16363 standard as well as its individual metrics and requirements
- ❑ Determine parameters, including which metrics do not apply to the repository
- ❑ Find a sponsor
- ❑ Determine a timeline for completing self-assessment
- ❑ Consider how to record findings
- ❑ Inform colleagues of this work and generate interest

Approach Stage 2

- ❑ Identify colleagues who are responsible for the work described in each metric
- ❑ Schedule individual meetings to discuss whether and how requirements are being met
- ❑ Record this information, as well as any relevant documentation, procedures, and standards

Approach Stage 3

(Results; internal wiki)

- ❑ Create main project area as well as a page for each metric
- ❑ Include the text for the metric and what evidence is required
- ❑ Record stakeholders interviewed and their feedback
- ❑ Embed or link to existing supporting evidence and record what's needed
- ❑ This also helps with transparency and maintaining an audit trail

Approach Stage 3 (Results; spreadsheet)

- Number and name of metric
- Short description of metric scope
- Supporting evidence
- System agnostic or system specific
- Current RAG status
- Action owner and/or actionee
- Timeframe and/or resource estimate

Organizational Infrastructure

3.3.5 The repository shall define, collect, track, and appropriately provide its information integrity measurements.

Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement

Written definition or specification of the repository's integrity measures (for example, computed checksum or hash value); documentation of the procedures and mechanisms for monitoring integrity measurements and for responding to results of integrity measurements that indicate digital content is at risk; an audit process for collecting, tracking, and presenting integrity measurements; Preservation Policy and workflow documentation.

Digital Object Management

4.4.1.2 The repository shall actively monitor the integrity of AIPs.

Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement

Fixity information (e.g., checksums) for each ingested digital object/AIP; logs of fixity checks; documentation of how AIPs and Fixity information are kept separate; documentation of how AIPs and accession registers are kept separate.

Infrastructure and Security Risk Management

5.1.1.3 The repository shall have effective mechanisms to detect bit corruption or loss.

Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement

Documents that specify bit error detection and correction mechanisms used; risk analysis; error reports; threat analysis; periodic analysis of the integrity of repository holdings.

Approach Stage 4

- ❑ Write a report of findings
- ❑ Circulate report to key organisational staff
- ❑ Determine next steps for implementing report recommendations
- ❑ Set a timeline for undertaking re-assessment

Approach (summary)


- Read the standard and make a plan
 - What is the scope?
 - How long will it take?
 - Who is my sponsor?
 - Who must I speak to?
 - How will I record my answers?
- Interview colleagues
- Obtain documentation
- Analyse and report to sponsor

Lessons Learned

- ❑ Work out what each metric is really asking you to avoid overlap
- ❑ Factor advocacy and comms into your plan to get stakeholder buy-in
- ❑ Don't rely on one interviewee to tell you everything
- ❑ Resource your assessment properly
- ❑ Document interviews and analysis in a way that's easy to transfer into your final report
- ❑ Understand how your recommendations can be implemented

“attaining trustworthy status is not a one-time accomplishment, achieved and forgotten. To retain trustworthy status, a repository will need to undertake a regular cycle of audit and/or certification.” – ISO 16363

Thank you!

Caylin Smith
Repository Manager
caylin.smith@bl.uk
 @caylinssmith