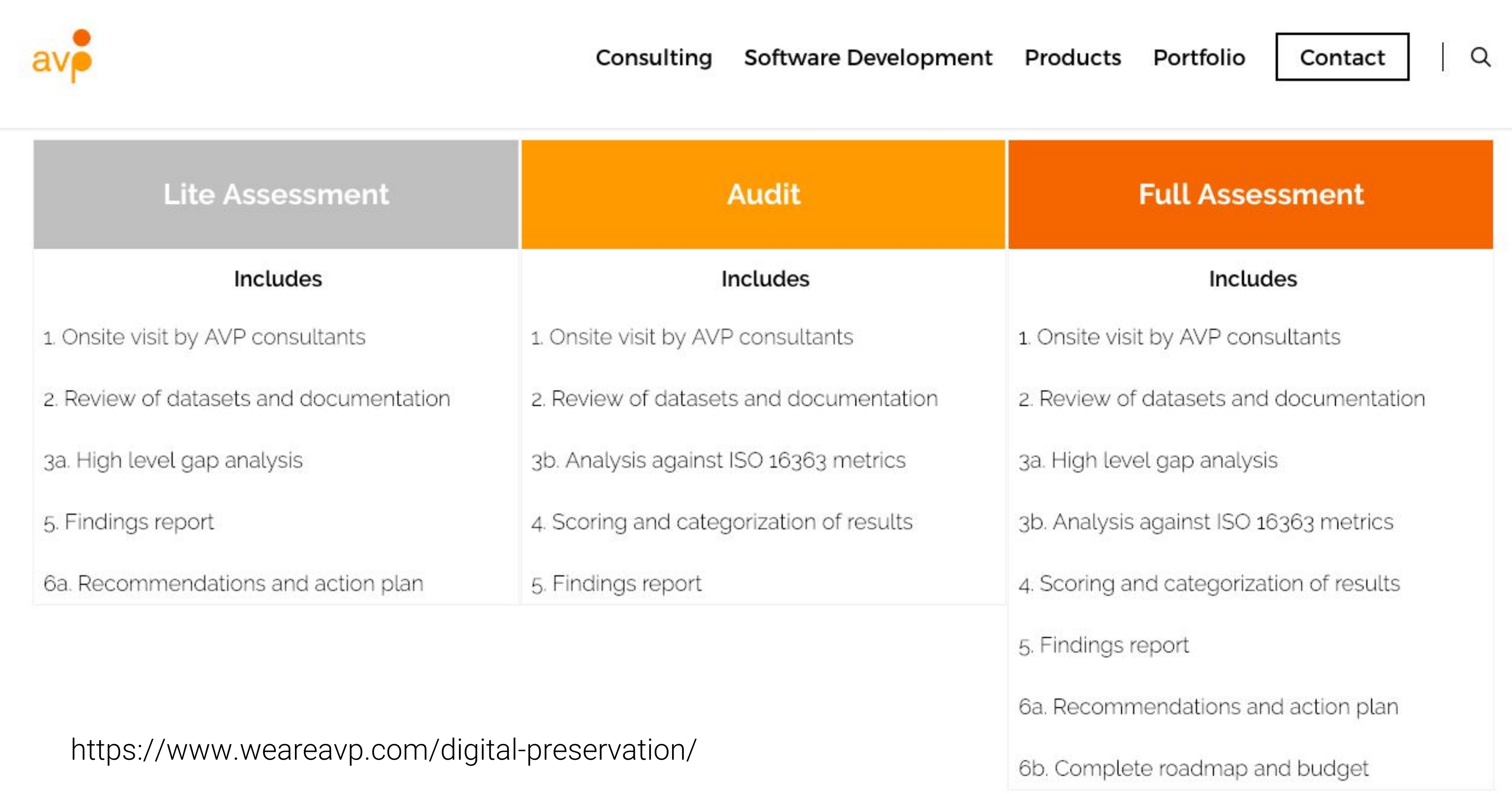
TOOLS & RESOURCES FOR SELF ASSESSMENT

Amy Rudersdorf
Senior Consultant, AVP

weareavp.com | @weareavp | amy@weareavp.com



Lite Assessment	Audit	Full Assessment
Includes	Includes	Includes
1. Onsite visit by AVP consultants	1. Onsite visit by AVP consultants	1. Onsite visit by AVP consultants
2. Review of datasets and documentation	2. Review of datasets and documentation	2. Review of datasets and documentation
3a. High level gap analysis	3b. Analysis against ISO 16363 metrics	3a. High level gap analysis
5. Findings report	4. Scoring and categorization of results	3b. Analysis against ISO 16363 metrics
6a. Recommendations and action plan	5. Findings report	4. Scoring and categorization of results
		5. Findings report
		6a. Recommendations and action plan
		6b. Complete roadmap and budget

https://www.weareavp.com/digital-preservation/

OUTLINE OF TALK

1. What is "assessment"?

2. What level of assessment do you want to perform?

3. Tools for assessment

   - Level 1: LoP
   - Level 2: CoreTrust Seal
   - Level 3: ISO 16363

4. How do you undertake assessment?

WHAT IS “ASSESSMENT”?

“the evaluation or estimation of the nature, quality, or ability of someone or something.”

DISTINCT FROM "AUDIT"

"an official inspection of an individual's or organization's programs, policies, and resources."

DISTINCT FROM “CERTIFICATION”

“the action or process of providing someone or something with an official document attesting to a status or level of achievement.”

WHICH LEVEL OF ASSESSMENT ARE YOU PREPARED TO PERFORM?

“least complex”

“moderate
”

“most complex”

LEAST COMPLEX

NDSA Levels of Preservation

https://ndsa.org/activities/levels-of-digital-preservation

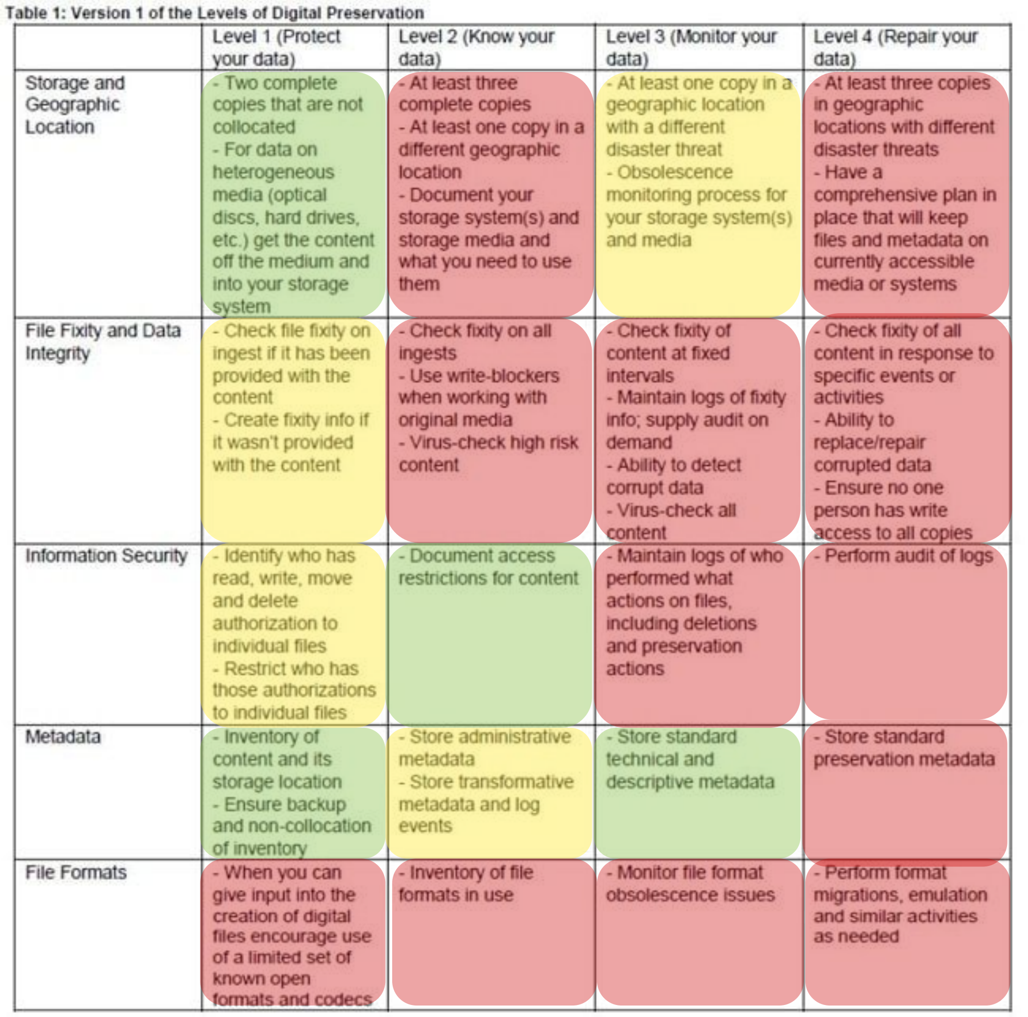


Table 1: Version 1 of the Levels of Digital Preservation

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	- Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system	- At least three complete copies - At least one copy in a different geographic location - Document your storage system(s) and storage media and what you need to use them	- At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media	- At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems
File Fixity and Data Integrity	- Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content	- Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content	- Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content	- Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	- Identify who has read, write, move and delete authorization to individual files - Restrict who has those authorizations to individual files	- Document access restrictions for content	- Maintain logs of who performed what actions on files, including deletions and preservation actions	- Perform audit of logs
Metadata	- Inventory of content and its storage location - Ensure backup and non-collocation of inventory	- Store administrative metadata - Store transformative metadata and log events	- Store standard technical and descriptive metadata	- Store standard preservation metadata
File Formats	- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs	- Inventory of file formats in use	- Monitor file format obsolescence issues	- Perform format migrations, emulation and similar activities as needed

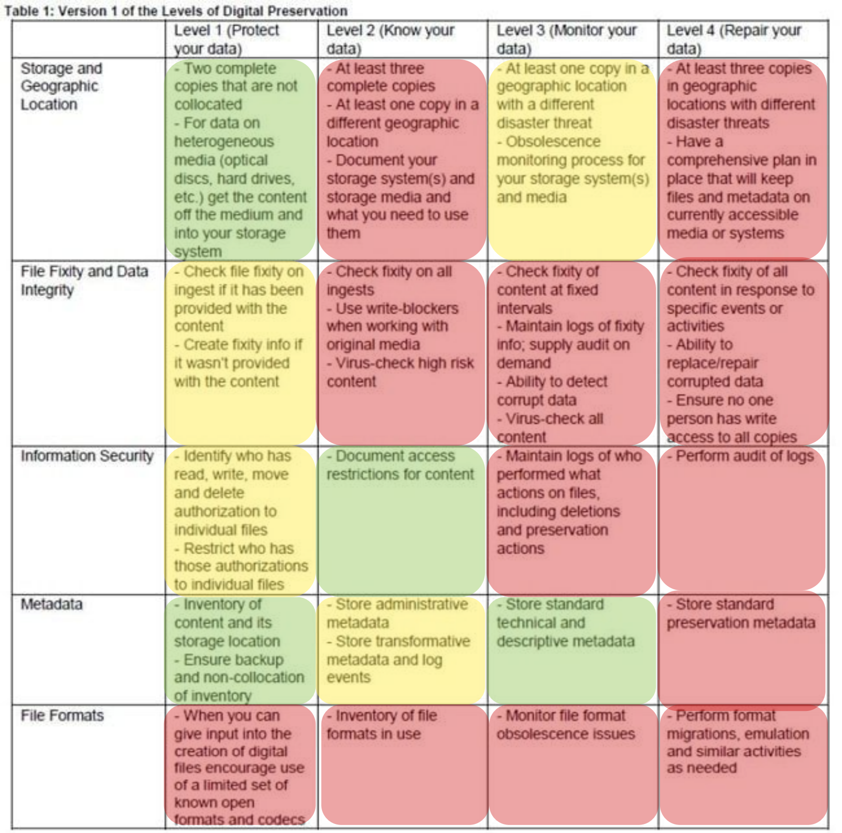


Table 1: Version 1 of the Levels of Digital Preservation

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	🟩 - Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system	🟥 - At least three complete copies - At least one copy in a different geographic location - Document your storage system(s) and storage media and what you need to use them	🟨 - At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media	🟥 - At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems
File Fixity and Data Integrity	🟨 - Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content	🟥 - Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content	🟥 - Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content	🟥 - Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	🟨 - Identify who has read, write, move and delete authorization to individual files - Restrict who has those authorizations to individual files	🟩 - Document access restrictions for content	🟥 - Maintain logs of who performed what actions on files, including deletions and preservation actions	🟥 - Perform audit of logs
Metadata	🟩 - Inventory of content and its storage location - Ensure backup and non-collocation of inventory	🟨 - Store administrative metadata - Store transformative metadata and log events	🟩 - Store standard technical and descriptive metadata	🟥 - Store standard preservation metadata
File Formats	🟥 - When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs	🟥 - Inventory of file formats in use	🟥 - Monitor file format obsolescence issues	🟥 - Perform format migrations, emulation and similar activities as needed

1.5/4

0.5/4

1.5/4

2.5/4

0/4

3/5 1.5/5 1.5/5 0/5

Overall score = 6/20

NDSA LoP REBOOT

Assessment subgroup is identifying ways that LoPs are being used for assessment (self, peer, vendor, etc.)

https://goo.gl/forms/AAJ04m41qE5Nobev2

MODERATE COMPLEXITY

CoreTrustSeal*

*Another “moderate” assessment tool is nestor.

CORE TRUST SEAL

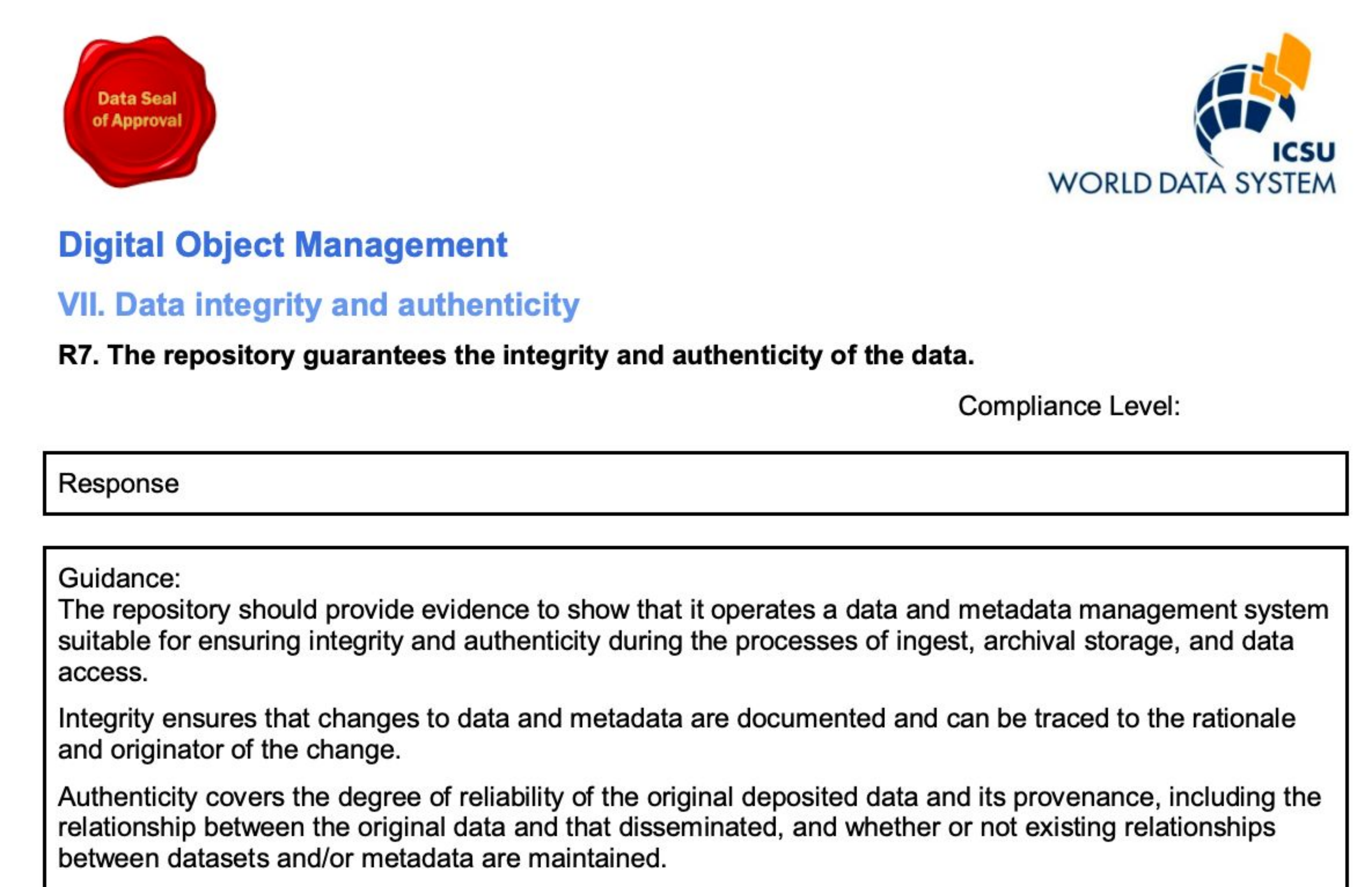
A core-level certification based on the DSA-WDS Core Trustworthy Data Repositories Requirements catalogue and procedures.

https://t2m.io/nmNri6EP

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

16 requirements / 3 sections

- Organizational Infrastructure
- Digital Object Management
- Technology



Data Seal of Approval

ICSU WORLD DATA SYSTEM

Digital Object Management

VII. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

Response

Guidance:
The repository should provide evidence to show that it operates a data and metadata management system suitable for ensuring integrity and authenticity during the processes of ingest, archival storage, and data access.

Integrity ensures that changes to data and metadata are documented and can be traced to the rationale and originator of the change.

Authenticity covers the degree of reliability of the original deposited data and its provenance, including the relationship between the original data and that disseminated, and whether or not existing relationships between datasets and/or metadata are maintained.

MOST COMPLEX

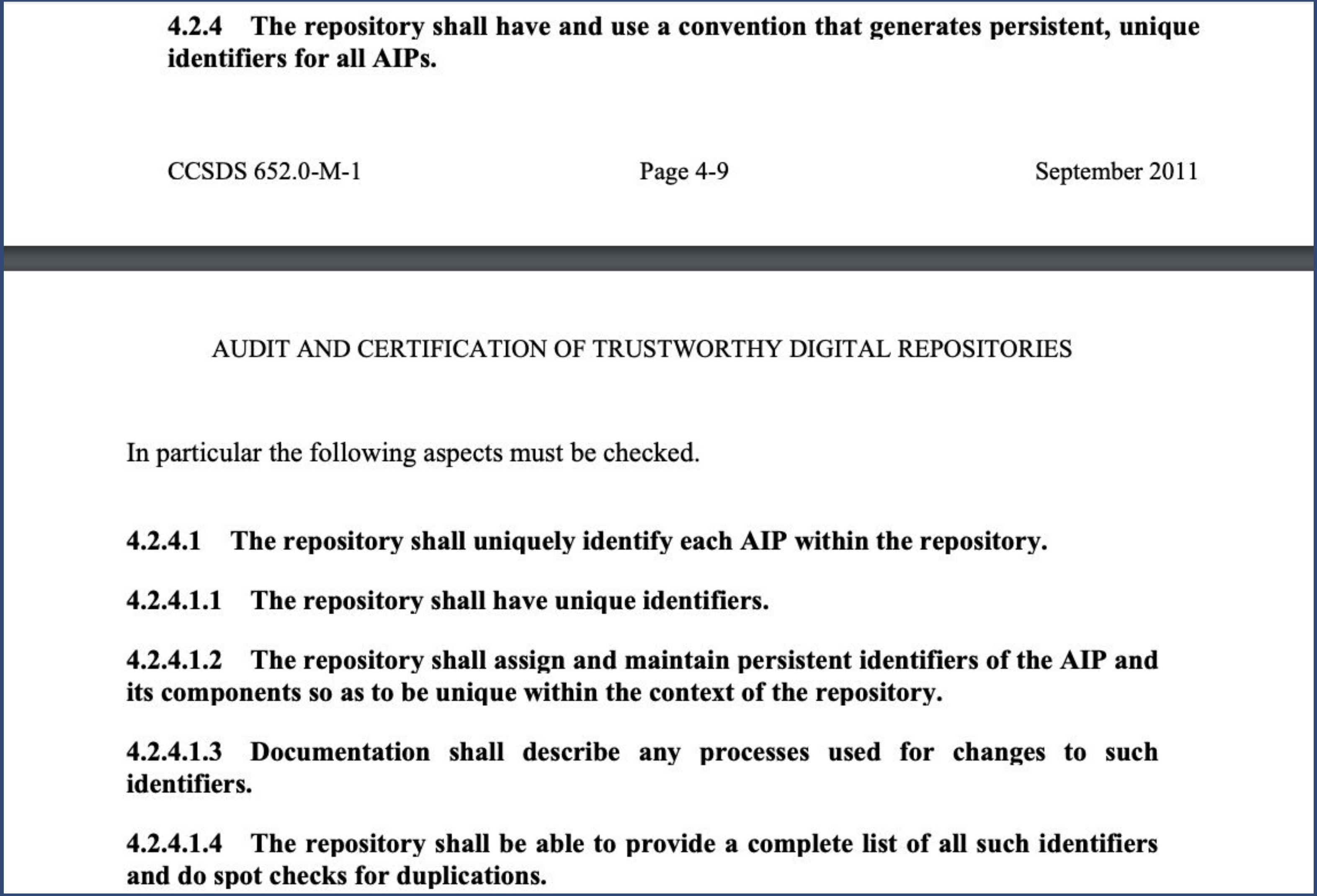
ISO 16363:2012 –
Audit and certification of trustworthy digital repositories

https://public.ccsds.org/pubs/652x0m1.pdf

Although “audit” and “certification” are in the name, AVP (and others) regularly use/s ISO 16363 for digital preservation assessments/self assessments.

3 sections / 109 metrics

- Organizational infrastructure
- Digital object management
- Infrastructure & security risk management



4.2.4 The repository shall have and use a convention that generates persistent, unique identifiers for all AIPs.

In particular the following aspects must be checked.

4.2.4.1 The repository shall uniquely identify each AIP within the repository.

4.2.4.1.1 The repository shall have unique identifiers.

4.2.4.1.2 The repository shall assign and maintain persistent identifiers of the AIP and its components so as to be unique within the context of the repository.

4.2.4.1.3 Documentation shall describe any processes used for changes to such identifiers.

4.2.4.1.4 The repository shall be able to provide a complete list of all such identifiers and do spot checks for duplications.

TIPS FOR SELF ASSESSMENT

Choose your assessment tool based on your organization's needs and capacity for "complexity."

*Anticipate spending several months on this process.

Meet with stakeholders to discuss each topic / metric.

Include admin, IT, creators / collections managers, dp staff, users.

Review extant policies, guidelines, & documentation.

** Note where these are missing or incomplete.

DOCUMENT EVERYTHING

Analyze results. Indicate:

Achievements

Strengths

Gaps

Challenges

Create a roadmap for addressing gaps.

(6 months, 1 year, 2 year, 3 year)

Identify roles & responsibilities for addressing gaps

DRAFT A FORMAL REPORT

- Stakeholders see what you are doing right!
- It's clear what is needed to support growth (staff, tech, $$)!
- Publicizing it makes it harder to ignore you / your team!

Schedule your next assessment for three years from date of completion.

OTHER TOOLS & SERVICES

AVP — we perform assessments!

weareavp.com/digital-preservation

NEDCC/Lyrasis digital preservation peer assessment

www.nedcc.org/preservation-training/digital-preservation-assessment-training

nestor

www.langzeitarchivierung.de/Subsites/nestor/DE/Siegel/siegel_node.html amy@WeAreAVP.com

WeAreAVP.com

@WeAreAVP