Performing Your Own Digital Preservation Assessment

Best Practices Exchange 2020

Amy Rudersdorf
Senior Consultant
@weareavp
amy@weareavp.com
https://weareavp.com
“[A]...flaw in the human character is that everybody wants to build and nobody wants to do maintenance.”

— Kurt Vonnegut
Performing A Digital Preservation Assessment

Outline

- Definition
- Why do it?
- Tools available?
- How to choose?
- What to expect?
- How to perform it?
- Using the outcomes
- Self, peer, vended?
What is...?
Digital Preservation Assessment; Self Assessment

Digital preservation assessments are performed to identify gaps and challenges (and successes!) that a repository faces.

Self Assessment - DIY for your repository!

The goal is to address those gaps to improve/implement processes, policies, governance, and workflows.
Where are we? (Status check)

Where do we want to go? (Goal planning)

How do we get there? (Roadmapping)

And ultimately, how do we get buy-in? (Communication)
**Why do it?**

*To address the gaps and challenges in current practice.*

<table>
<thead>
<tr>
<th>To understand where your org is today.</th>
<th>To understand where you want to go.</th>
<th>To understand how you get there.</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the current status of your repository?</td>
<td>Set goals.</td>
<td>What resources will it take to achieve goals in box #2?</td>
</tr>
<tr>
<td>- Gaps</td>
<td>Align with best practices/standards.</td>
<td>How do you prioritize?</td>
</tr>
<tr>
<td>- Challenges</td>
<td>Identify internal factors influencing decision making.</td>
<td>How long will it take?</td>
</tr>
<tr>
<td>- Successes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## What are your options?

**Digital Preservation Assessment Tools landscape (sampling)**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Managed by?</th>
<th>Area of Focus</th>
<th>Cost to Use</th>
<th>Complexity</th>
<th>Certification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CoreTrustSeal</strong></td>
<td>CoreTrustSeal organization</td>
<td>Data</td>
<td>Free to use; €1,000 for certification</td>
<td>Low/Medium</td>
<td>Yes, but not required</td>
</tr>
<tr>
<td><strong>ISO 16363 (TDR)</strong></td>
<td>International Standards Organization (ISO)</td>
<td>Holistic audit of org as relates to the OAIS framework</td>
<td>Must purchase the standard (~$180) from ISO</td>
<td>High</td>
<td>Yes, but not required</td>
</tr>
<tr>
<td><strong>Levels of Preservation</strong></td>
<td>NDSA</td>
<td>Technology and metadata focused</td>
<td>Free</td>
<td>Low/Medium</td>
<td>No</td>
</tr>
<tr>
<td><strong>NEDCC Peer Assessment</strong></td>
<td>Northeast Document Conservation Center</td>
<td>Holistic; aligns with 16363</td>
<td>Free</td>
<td>Low/Medium</td>
<td>No</td>
</tr>
<tr>
<td><strong>Rapid Assessment Model</strong></td>
<td>Digital Preservation Coalition (DPC)</td>
<td>Holistic; aligns with 16363</td>
<td>Free</td>
<td>Low/Medium</td>
<td>No</td>
</tr>
<tr>
<td><strong>Vendor assessment</strong></td>
<td>Vendor</td>
<td>Aligned with institutional need</td>
<td>Vendor fees</td>
<td>Low</td>
<td>Maybe</td>
</tr>
</tbody>
</table>
What to Expect?
Digital Preservation Assessment

- **Resource intensive**, and not only for digital preservation staff (consider other stakeholders, such as administration and IT)
- **Deep analysis** of governance models, finances, legal commitments, designated communities, in addition to technology
- **Honest discussions** about challenges and gaps
- **Documenting outcomes** for a wide audience (depending on goals)
What will you do during the assessment?

Digital Preservation Assessment

- Documentation review
- Discovery
- Gap analysis
- Recommendations
- Reporting
- Roadmapping
How to perform a self assessment?
Step-by-step process

1. Choose a tool
2. Gather documentation
3. Perform discovery
4. Analyse results
5. Document results
6. Build a roadmap
# Select a Tool

Digital Preservation Assessment Tools landscape (sampling)

<table>
<thead>
<tr>
<th>Tool</th>
<th>Managed by?</th>
<th>Area of Focus</th>
<th>Cost to Use</th>
<th>Complexity</th>
<th>Certification</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoreTrustSeal</td>
<td>CoreTrustSeal organization</td>
<td>Data</td>
<td>Free to use; €1,000 for certification</td>
<td>Low/Medium</td>
<td>Yes, but not required</td>
</tr>
<tr>
<td>ISO 16363 (TDR)</td>
<td>International Standards Organization (ISO)</td>
<td>Holistic audit of org as relates to the OAIS framework</td>
<td>Must purchase the standard (~$180) from ISO</td>
<td>High</td>
<td>Yes, but not required</td>
</tr>
<tr>
<td>Levels of Preservation</td>
<td>NDSA</td>
<td>Technology and metadata focused</td>
<td>Free</td>
<td>Low/Medium</td>
<td>No</td>
</tr>
<tr>
<td>NEDCC Peer Assessment</td>
<td>Northeast Document Conservation Center</td>
<td>Holistic; aligns with 16363</td>
<td>Free</td>
<td>Low/Medium</td>
<td>No</td>
</tr>
<tr>
<td>Rapid Assessment Model</td>
<td>Digital Preservation Coalition (DPC)</td>
<td>Holistic; aligns with 16363</td>
<td>Free</td>
<td>Low/Medium</td>
<td>No</td>
</tr>
<tr>
<td>Vendor assessment</td>
<td>Vendor</td>
<td>Aligned with institutional need</td>
<td>Vendor fees</td>
<td>Low</td>
<td>Maybe</td>
</tr>
</tbody>
</table>
DPC Rapid Assessment Model

What is DPC RAM?

The DPC Rapid Assessment Model (RAM) is a digital preservation maturity modelling tool that has been designed to enable rapid benchmarking of an organization's digital preservation capability.

This model aims to be:

- Applicable for organizations of any size and in any sector
- Applicable for all content of long-term value
- Preservation strategy and solution agnostic
- Based on existing good practice
- Simple to understand and quick to apply

DPC Members, login to watch the introductory webinar on the DPC RAM

Where is DPC RAM?

The model is freely available to all. DPC Members login to the website to access online data entry form and additional member benefits.

- **The DPC Rapid Assessment Model** - the full model, introductory text and a worksheet
- **Digital worksheet** - an excel worksheet to record your assessment results

DPC Members login to see the online form for submitting your results
### ISO 16363 Alignment

Three main sections represented in RAM

1. **Organizational infrastructure**
   
   A, B, C, E, F

2. **Digital Object Management**
   
   G, H, I, J, K

3. **Infrastructure and Security Risk Management**
   
   D
This assessment tool comes with a spreadsheet in which you can organize and document your research results.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DPC RAM</td>
<td>Documenting results</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Digital Preservation Coalition Rapid Assessment Model (DPC RAM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Assessment Worksheet</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Organization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Assessment completed by</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Assessment complete on</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Note on scope of assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Current Level</th>
<th>Evidence/Notes</th>
<th>Target Level</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Operational viability</td>
<td>Governance, organizational structure, staffing and resourcing of digital preservation activities.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Policy and strategy</td>
<td>Policies, strategies, and procedures which govern the operations and management of the digital archive.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Legal basis</td>
<td>Management of contractual, licensing, and other legal rights and responsibilities relating to acquiring, preserving and providing access to digital content (e.g., licensing, copyright, terms and conditions of use, data protection regulations).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>IT capability</td>
<td>Information Technology capabilities for supporting digital preservation activities.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Continuous improvement</td>
<td>Processes for the assessment of current digital preservation capabilities, the definition of goals and the monitoring of progress.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>Community engagement</td>
<td>Engagement with and contribution to the wider digital preservation community.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Acquisitional and ingest</td>
<td>Processes to acquire or transfer content and ingest it into a digital archive.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>Minimum preservation</td>
<td>(Provision to ensure...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How to perform a self assessment?

Step-by-step process

Choose a tool
Gather documentation
Perform discovery
Analyse results
Document results
Build a roadmap
Gather documentation
Supplement the discovery interviews you perform and support findings.

- Policies & planning documents
- Workflows, process docs, & infrastructure/tech diagrams
- Program reports & directives
- Org charts, staffing & resource docs
How to perform a self assessment?

Step-by-step process

1. Choose a tool
2. Gather documentation
3. Perform discovery
4. Analyse results
5. Document results
6. Build a roadmap
“Listening is not understanding the words of the question asked, listening is understanding why the question was asked in the first place.”

— Simon Sinek
What are the guiding principles for current and future digital collections and digital preservation systems?

What is the vision for how digital collections are managed within the organization? This might relate to resources, staffing, technology, access, etc.

From your vantage point, what are the pain points for digital collections management and organization? Where are the successes?

What type of fiscal and organizational support is there for digital collections and digital preservation now and moving forward?

Support for staffing growth and/or professional development?

What is your primary responsibility to the administration/management of technology supporting the creation, collection, use, and/or preservation of digital collections?

Talk about hosting and storage policies, including redundancy, geographic distribution of storage systems, and security. What types of storage are in place (spinning disk, LTO, cloud, etc.)? What is their current capacity? Are there plans for growth? Are there plans for adopting a new/different approach?

How much storage is required now? Future growth in # of files and size in TB/PB?

Do your storage devices allow for dynamic fixity checks?
How to perform a self assessment?

Step-by-step process

1. Choose a tool
2. Gather documentation
3. Perform discovery
4. Analyse results
5. Document results
6. Build a roadmap
Analysing discovery interviews and documentation

Identifying gaps, challenges, and successes

1. Color code & map
Read through and mark up your documentation and discovery interview info first. Color code areas relevant to the RAM model. For example, references to metadata are blue; storage are yellow. Now you can map your findings to the RAM categories (or other model).

2. De-dupe & synthesize
You’ve mapped all of the relevant findings. There’ll be duplicates where two people say the same thing, for example. De-dupe and synthesize those responses. You don’t need long narratives; keep verbosity to a minimum.

3. Identify gaps
Once you’ve synthesized all findings and mapped them to your tool, you can score your results to identify gaps and successes. Your recommendations will be based on the gaps you have identified.
How to perform a self assessment?

Step-by-step process

1. Choose a tool
2. Gather documentation
3. Perform discovery
4. Analyse results
5. Document results
6. Build a roadmap
Documenting results

Who is your audience/what is your goal?

Note: Document findings in DPC RAM spreadsheet for digital staff.

5.2 Gaps and recommendations

ISO 16363: Audit and certification of trustworthy digital repositories (TDR) is a standard for assessing the trustworthiness of digital repositories. For purposes of consistency, "Repository" is used in this report, as defined by ISO 16363, to mean "the organization responsible for digital preservation."

In this way, it is more expansive than the definition of repository, which currently includes description, storage, and management (for limited formats) using the applications, infrastructure, and workflows maintained by the IT and IT units. The rationale of TDR, and the reason we are invoking it for this report, is that a critical component of digital archiving infrastructure is the ability to support description, storage, migration, and access to digital collections. In addition, TDR stresses the need for broader organizational support and oversight, object management that ensures the sustainability of and access to the assets held by an organization, and the infrastructure to store and maintain those objects over time.

In regard to the gaps and recommendations listed below, we suggest taking a step back and considering all aspects of the repository as they are described by TDR. TDR describes the three main categories of a digital repository as a three-legged stool, where each category is a leg that holds up the stool. Without one of the categories (or legs), the stool will not stand. The three categories in TDR we will use to frame our findings are:

- **Organizational Infrastructure**
  Although "adequate technical architecture, processes, and capabilities underpin a trusted digital repository, the technical aspects are only one piece of the overarching infrastructure supporting the digital repository functions. Organizational attributes of digital repositories are equally critical. Organizational attributes are characteristics of the repository organization that affect performance, accountability, and sustainability."
  These can be organized into four attributes: administrative responsibility, organizational and administrative coordination, technical standards, and digital infrastructure.

! Note: Other stakeholders may need a more narrative report in which you clearly state the challenges and your recommended responses.

Reminder: To gain buy-in, document your findings in a way that communicates best to your decision makers.
How to perform a self assessment?

Step-by-step process

Choose a tool
Gather documentation
Perform discovery
Analyse results
Document results
Build a roadmap
“There's an old Wayne Gretzky quote that I love. ‘I skate to where the puck is going to be, not to where it has been.' And we've always tried to do that at Apple.”

— Steve Jobs, 2010
6. Roadmap

AVP’s proposed sequencing and prioritization of the recommendations in this report are listed below. For a more robust description of each recommendation, please refer to section 5.2 of this report.

<table>
<thead>
<tr>
<th>Action Items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Priority 1 (0-6 months)</strong></td>
</tr>
<tr>
<td>Establish a dedicated Digital Preservation Advisory Group (5.2.1.1, R1)</td>
</tr>
<tr>
<td>Hire or assign a dedicated IT systems position to support the digital preservation software and infrastructure (5.2.1.4, R9)</td>
</tr>
<tr>
<td>Select and implement a digital preservation system (5.2.3.2, R14)</td>
</tr>
<tr>
<td>Determine infrastructure requirements to support long-term storage of digital assets (5.2.3.3, R16)</td>
</tr>
<tr>
<td><strong>Priority 2 (6-12 months)</strong></td>
</tr>
<tr>
<td>Hire a dedicated Director of Digital Preservation position to lead the digital preservation team at NYPL (5.2.1.4, R7)</td>
</tr>
<tr>
<td>Hire a dedicated Digital Repository Manager/Digital Archivist to serve as the business manager for the new digital preservation system (5.2.1.4, R8)</td>
</tr>
<tr>
<td>Review, update, approve, and distribute current digital preservation policy and planning documentation to guide development of the program (5.2.1.2, R3)</td>
</tr>
<tr>
<td>Harness the Digital Preservation Advisory Group to identify, prioritize, document, and address policy gaps that impede consistent digital preservation workflows and practices (5.2.1.2, R2)</td>
</tr>
<tr>
<td>Harness the Digital Preservation Advisory Group to prioritize and generate documentation around existing systems from each of the areas that interact with the digital preservation system (5.2.1.3, R4)</td>
</tr>
</tbody>
</table>

**DON’T SKIP THIS STEP!**

- Priority 1 (0-6 months)
- Priority 2 (6-12 months)
- Priority 3 (12-24 months)
- Priority 4 (24-36 months)

**Recommendation:** always include a new assessment activity in the Priority 4 section.
How to perform a self assessment?

Step-by-step process

Choose a tool
Gather documentation
Perform discovery
Analyse results
Document results
Build a roadmap
What do you do with your findings?

Share!

For Digital Staff
- Acknowledge gaps
- Use recommendations to address gaps
- Use roadmap to prioritize your work

For Admin
- Use report to communicate
  - value of program
  - plan for improvement
  - need for policies
- Use report to push for formal programmatic support.

For IT
- Use report to communicate
  - infrastructure gaps
  - storage needs
  - required dp actions
- Use report to push for formal programmatic support.

For Other Stakeholders
- Use report to communicate
  - program successes
  - plan for growth and programmatic improvements
- Use report to develop trust.

Share!
What are your options?
Digital Preservation Assessment

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Staffing Required</th>
<th>Time Required</th>
<th>$$ Required</th>
<th>Your knowledge of DP concepts required</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIY</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Peer + DIY</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Vended</td>
<td>Low/Medium</td>
<td>Low</td>
<td>Medium/High</td>
<td>Low/Medium</td>
</tr>
</tbody>
</table>

Consider:
What questions need answers?
Is tool complexity OK?
Do you have the skilled staff?
Does that staff have time?
How much $$ can you spend?
Institutional capacity?
“Our culture's obsession with innovation and hype has led us to neglect maintenance and maintainers. . . . There's not only a need, but a certain nobility in taking care of what you've already created, and maybe we shouldn't look at maintenance as the enemy of innovation.”

AVP is a data management consulting and software firm focused on freeing organizations from the obstacles of information management.

QUESTIONS?

Amy Rudersdorf
Senior Consultant
@weareavp
amy@weareavp.com
https://weareavp.com