Final Report of the Survey on Digital Preservation Capacity and Needs at Canadian Memory Institutions, 2017-18

Produced by **Grant Hurley** and **Kathleen Shearer**
for the CARL Digital Preservation Working Group

NOVEMBER 2019

Corresponding dataset DOI: 10.5683/SP2/8KLEPP

www.carl-abrc.ca

CARL ABRC
CANADIAN ASSOCIATION OF RESEARCH LIBRARIES
ASSOCIATION DES BIBLIOTHÈQUES DE RECHERCHE DU CANADA

Table of Contents

Executive Summary

Digital preservation presents one of the greatest challenges for memory and cultural institutions in the modern age. New resources and specialized knowledge must match the complex organizational, policy and technical components that can assist in the persistence of digital content into the future. How can we ensure that the digital content of value to Canadians is available in perpetuity? What are the gaps that need to be addressed, and how can Canadian institutions be supported to build capacity in this area? In 2017 and 2018, the Canadian Association of Research Libraries (CARL) surveyed libraries, museums, galleries and archives to determine the current state of digital preservation in Canada and to better understand the issues and needs stemming from this work.

This report presents a summary of the results of the CARL Digital Preservation Working Group (DPWG) survey on digital preservation capacity and needs at Canadian memory institutions. The purpose of the survey was to provide an updated and comprehensive picture of digital preservation activities in Canada and to identify existing gaps and outstanding needs at Canadian institutions. Phase 1 of the survey targeted the 29 CARL members in October-December 2017 (with an update period in December 2018-January 2019) and Phase 2 targeted other Canadian memory institutions in August-September 2018. A total of 52 complete responses were received consisting of 27 research libraries and government-based CARL members, 7 academic libraries (not CARL members); 13 government-based archives, libraries, and museums at the municipal, provincial and national levels; and 5 community-based or non-profit archives, libraries and museums. Of the government-based respondents, Library and Archives Canada responded as a member of CARL in Phase 1, and 5 provincial and territorial archives responded in Phase 2.

All respondent organizations are undertaking **digitization** activities, and 94% are **collecting** born-digital materials. The main sources of born-digital materials are records and publications from the respondent organizations themselves and personal records from private donors, followed by content produced by faculty members and researchers. 3 organizations are not collecting born-digital materials due to a lack of capacity to preserve them.

Organizational commitment to digital preservation is in development among respondents. 38% of organizations have language that expresses a commitment to digital preservation published in a strategic plan, mandate or mission statement. Another 40% have language waiting approval, in draft, or planned. 81% of organizations have an individual or group responsible for coordinating digital preservation activities across the organization and 44% have a committee or working group related to digital preservation work. Engagement with external organizations,

projects and initiatives related to digital preservation is high, with 75% of respondents indicating participation in these groups.

The availability of formalized digital preservation **policies and procedures** is low among respondents, but progress is being made in this area as many organizations begin to scope and draft them. Only 17% of respondent organizations have an approved digital preservation policy or set of policies, though an additional 19% have policies under review or in draft. In contrast, more organizations are working to document preservation procedures, strategies and plans. 23% have documented procedures, and another 48% have procedures in draft or development. 67% have digital preservation strategies or plans in place. 58% of organizations also indicated they have adopted specific digital preservation-related standards, best practices or guidelines.

The use of **tools for digital preservation** tasks among respondents is low. 29% of respondent organizations are using digital forensics tools to safely capture data from storage media, and 48% are using at least one tool for preservation processing. However, many of these uses of tools are still in a testing period: less than half of the respondents using tools for processing are using any one tool in production.

Most respondents endeavour to give **access** to digital materials. 85% of respondent organizations use a variety of web-based platforms for access, and others use shared folders or computers on-site.

The transition to preservation-friendly **storage** among respondent organizations has been slow. While 94% of organizations rely on local network storage as one storage option, 81% also have digital assets stored on CDs/DVDs, hard and flash drives or legacy media (such as floppy disks), and 38% selected all three options. 50% of respondent organizations make use of cloud storage services. 50% use tape and 35% make use of replicated storage networks like LOCKSS. Respondents are storing a median of 20TB of digital content.

There are low **staffing** levels devoted to digital preservation at many organizations. Looking at total Full Time Equivalent (FTE) values across respondent organizations, positions with responsibilities for digital preservation represent less than 1% of total organizational FTEs. 62% of respondents have less than the equivalent of one full-time individual working on digital preservation across all staff listed with responsibilities in this area. Responsibilities are often spread across several positions, with 48% of all roles listed by respondents having between 0 and 20% of a person's time given to digital preservation. Expectations for expanding staffing are mixed: 48% said they intended to increase staffing through new hires or reassignment.

Digital preservation programs are largely **funded** through general operational budgets, but 42% of respondent organizations also rely on short-term funding sources such as grants or awards to accomplish this work. 75% of respondents did not know what percentage of their organization's budget is dedicated to digital preservation. Outside of reliance on IT departments, few organizations rely on additional resources from elsewhere within their organization or externally.

Based on a scoring method used across quantitative responses in the survey, the following picture emerges **overall**. A small cohort of 13% of respondent organizations are just starting out in developing digital preservation capacity. A large middle group of 79% of respondent organizations have programs that are in development. Another small cohort of 8% of respondents showed strong capacity overall as their programs are based on more mature policy platforms, a stronger organizational commitment and improved funding devoted towards digital preservation.

In conclusion, Canada's cultural, research and educational content is an invaluable resource for our country and it is critical that we collectively ensure that this content is available to Canadians now and in the future. The results of the survey show that there are still many challenges for Canada's memory institutions related to digital preservation. In particular, memory organizations, many of which have mission or vision statements related to preservation and long-term access, have not yet prioritized digital preservation within their operations and do not direct significant resources towards these activities. The survey shows that many memory institutions, both large and small and across different sectors, are struggling with similar challenges related to digital preservation, and collective action in Canada could be helpful in addressing some of the issues. Training and skills development covering a range of aspects related to digital preservation (e.g. policy, organization, and technology) that are accessible to all the memory institutions in Canada would be beneficial. Additionally, organizational models will need to change to reflect new priorities related to digital preservation. Developing best practice and modern organizational models in the digital age that more appropriately reflect digital preservation needs can provide the impetus for institutions to begin the restructuring process that is required.

A. Introduction

This report presents the results of the CARL Digital Preservation Working Group (DPWG)'s 2017-18 survey of Canadian memory institutions on the state of digital preservation activities and needs at their organizations. The purpose of the CARL DPWG survey was to provide an updated and comprehensive picture of current digital preservation activities in Canada and to help identify existing gaps and outstanding needs at Canadian institutions. The report summarizes the results of the 2-phase survey in several areas: scope and type of digital content collected, organizational commitment, policies and procedures, tools and applications, discovery and access, storage, and staffing and funding. The report also provides an analysis of the results in terms of gaps, strengths and weaknesses. The appendices contain details of the comparison methodology and the complete list of survey questions, as well as more detailed responses for certain questions. It is anticipated that the findings will inform the development of strategies, policies, expertise and resource allocation and, we hope, lead to increased capacity and opportunities for collaboration and coordination in Canada around digital preservation. In addition, the survey results could serve as a baseline for measuring the progress related to digital preservation activities in Canadian memory institutions.

B. Methodology

The survey was developed by the authors in conjunction with the Survey Subgroup of the CARL Digital Preservation Working Group. The complete list of survey questions and sources consulted during the development of the survey are listed in Appendix 1. Respondents could fill out the survey in English or French. The survey consisted of 66 questions inputted in SimpleSurvey, a hosted software tool used by CARL. Some areas for improving the survey are noted in Appendix 6.

The survey was undertaken in two phases. Phase 1 surveyed members of CARL and was conducted between October and December 2017. It was distributed via the CARL Directors to relevant staff at member institutions. 26 CARL members provided complete responses.

Phase 2 targeted a wider set of memory institutions in Canada and was conducted between August and September 2018. The second phase was undertaken to broaden survey responses beyond research-driven academic libraries, and to understand whether these stakeholders have similar or different challenges related to digital preservation. Phase 2 was sent to the listservs Archives Canada listserv (Arcan-l), CanLib-Data (Research data management), and the CRKN Directors list, as well as to representatives of provincial and territorial archives via the Canadian Council of

Archives, and to executive directors of the Canadian Museums Association and the Ontario Association of Art Galleries for distribution to members. 25 organizations provided complete responses in Phase 2. They consisted of 7 academic libraries and archives outside of CARL; 13 government-based archives, libraries, and museums at the municipal, provincial and national levels; and 5 community-based or non-profit archives, libraries and museums.

6 CARL members (including one additional CARL member who did not participate in Phase 1) also submitted new responses as part of Phase 2. This prompted CARL to contact the remaining CARL members who responded in Phase 1 and invite them to revise and update their responses. 19 CARL members opted to do so during December 2018 and January 2019, and these responses were merged into the dataset to make for a final dataset consisting of 27 CARL members and 25 additional Canadian memory institutions for a total of 52 complete responses.

C. Dissemination

In addition to this report, the results of both phases were presented at the @Risk North 2: Digital Collections forum in Montréal on November 9, 2018. Summary reports were written for both Phase 1 and Phase 2 and published as part of the @Risk North 2 materials and can be consulted for more granular details specific to each phase. However, these reports do not include the 2018-19 updates from CARL respondents that are analyzed as part of the final dataset in this report.

The final dataset of responses and accompanying codebook is available in the Scholars Portal Dataverse: https://doi.org/10.5683/SP2/8KLEPP. Potentially identifying information was removed from the public version. The dataset includes 52 complete responses (marked as "Submitted") in addition to 8 incomplete responses (marked as "In progress").

D. Results

Notes:

- *Unless otherwise indicated, statistics are given using n=52 as the sample size.*
- *For each of the "gaps and challenges" sections, please note that participants could select all options that applied.*
- *Summary results are provided across all respondents. Those interested in more detailed sector-specific information (for example, comparisons between government-based respondents and community/non-profit respondents) should consult the first and second phase reports.*

- *Question numbers are indicated in brackets in the discussions of the results. Please refer to Appendix 1 for the corresponding question text.*

1. Respondent Profiles

The 52 respondents break down into the following sectors (Q4):

- 27 (52%) CARL members (including 1 government-based CARL member)
- 7 (13%) from academic libraries that are not members of CARL
    - 33 respondents in total are from the academic sector (63%).
- 13 government-based organizations (25%)
- 5 community-based or non-profit organizations (10%).

Respondents were asked to indicate organization subtypes and could select more than one response (Q5):

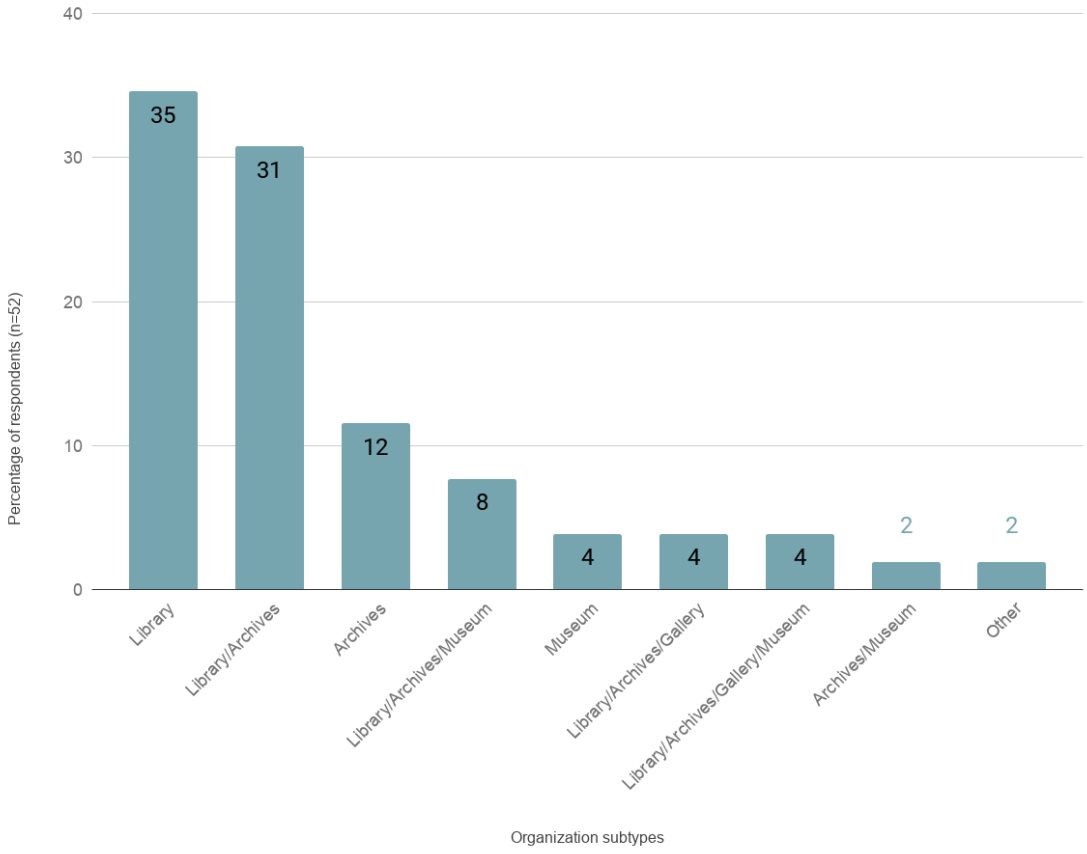

Figure 1: Combinations of organizational subtypes, ordered by frequency (Q5).

The total organizational budgets of respondents were self-reported as follows (Q6):

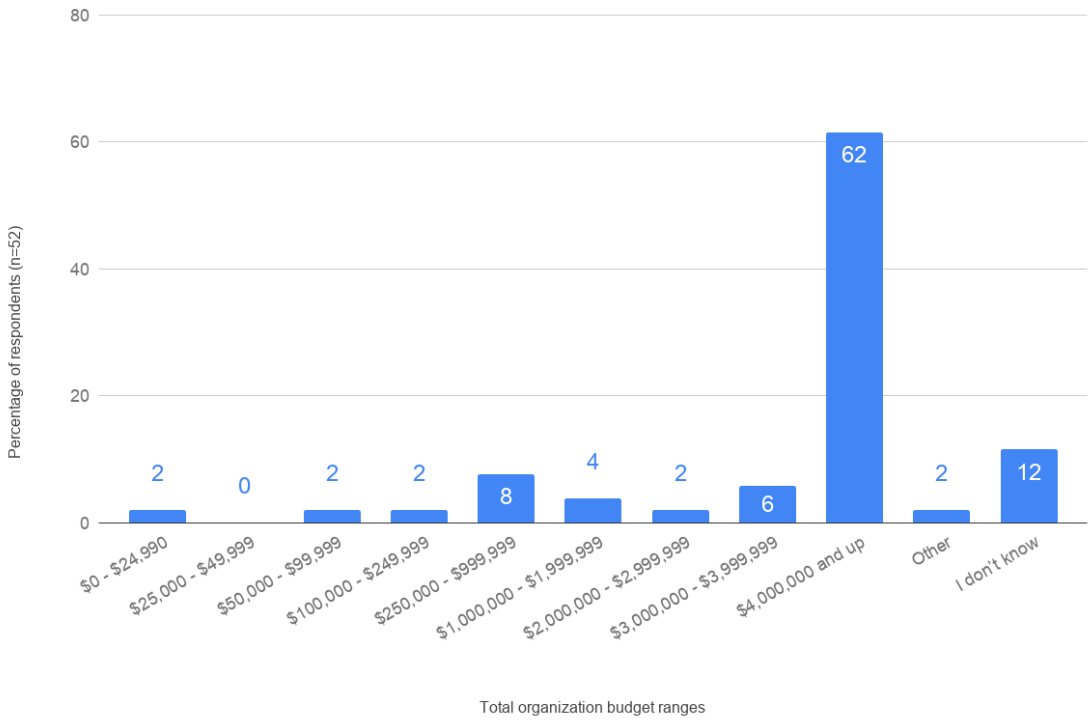

Figure 2: Total organization budget ranges as indicated by respondents (Q6)

Phase 2 respondents were asked to report on total yearly materials expenditures in order to complement official CARL statistics on these numbers (Q7). See Section E.2 for analysis using these figures. 9 respondents gave a figure. The median was $544,000, with a low of $30,000 and a high of $4,803,000.

Phase 2 respondents were also asked to report on total staffing levels as full time equivalents (FTEs) in their organizations (Q8). 22 provided a response. The median was 34.25 FTE, with a low of 1 FTE and a high of 624 FTE.

2. Scope and Types of Digital Collections

94% of respondents indicated that they are collecting born-digital materials (Q53) and **100% are involved in digitization activities** (Q56). Therefore, all respondent organizations have an interest in preserving these digital assets. The 3 respondents that said they are not collecting born-digital materials were government units: a public library, an archive, and an information management unit, all of whom indicated

that they are not collecting digital materials because they do not yet have the capacity to preserve them. For the respondents who are collecting born-digital materials, the materials came from two major sources: internal sources such as institutional records and publications, and private donors (Q54). Unsurprisingly, faculty and students represented a key source of content for the academic respondents, while government-sourced materials were a focus for both government-based and academic respondents.

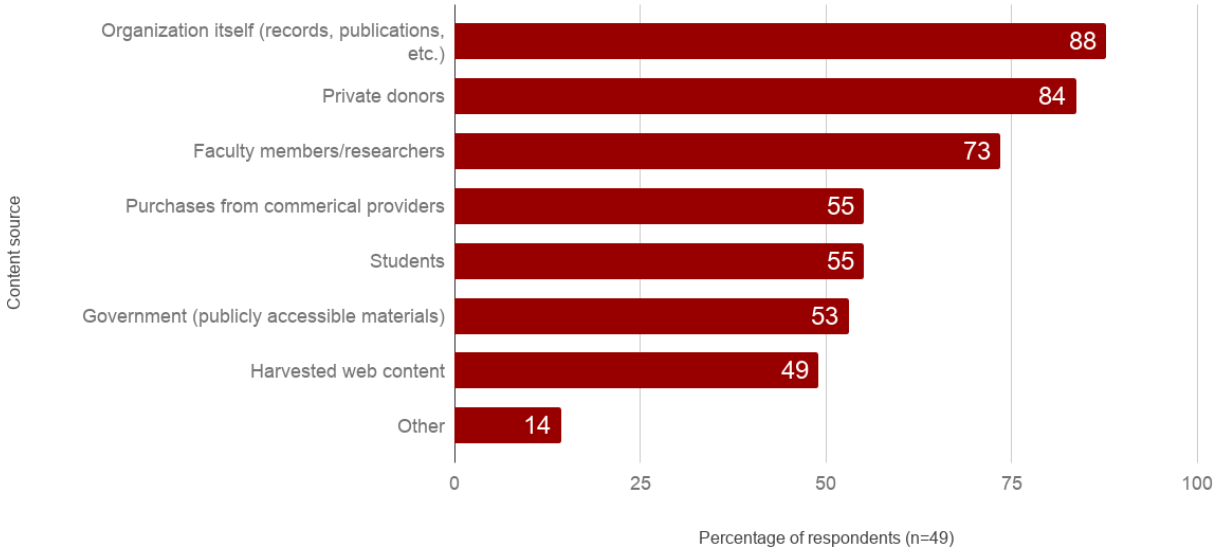

Figure 3: Sources of digital content collected ranked by percentage of respondents (Q54). Percentages were calculated out of the 49 respondents who indicated they are collecting born-digital materials.

The content types collected show alignment with the content sources (Q55). Materials commonly acquired by archives (photographs, moving images, audio, and documents in the form of personal papers and records) ranked the highest. The materials typically collected by libraries, such as purchased digital books, serials and datasets, roughly matched the range of respondents from the academic sector, though other types of organizations like government-based libraries also collect these kinds of materials.

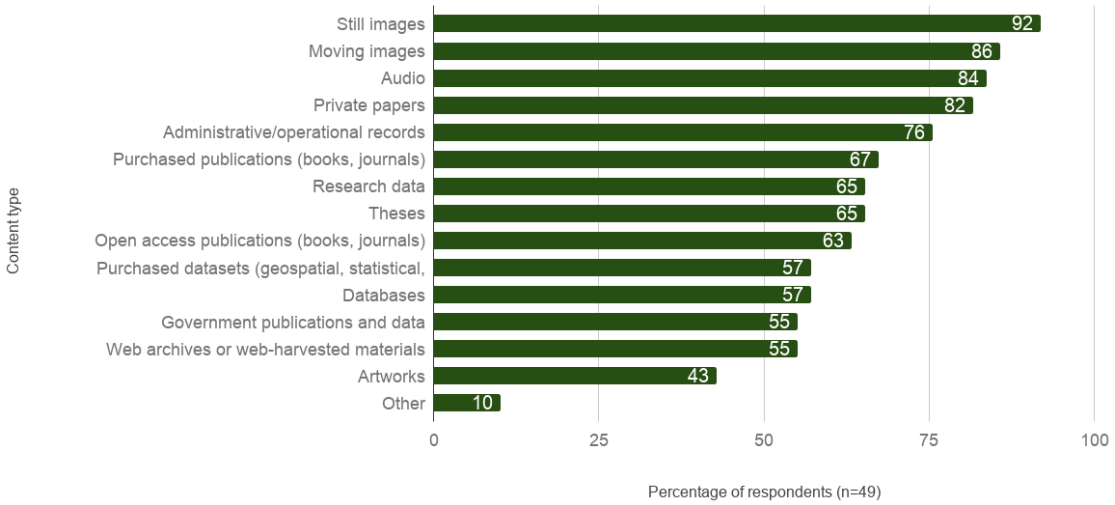

Figure 4: Types of digital content collected ranked by percentage of respondents (Q55). Percentages were calculated out of the 49 respondents who indicated they are collecting born-digital materials.

Respondents were asked to rank the **highest priorities for digitization** activities from 1 (lowest priority) to 5 (highest priority) (Q57). Of the 39 respondents (75%) who gave full answers, they ranked the materials in the following order:

- Photographs (4.13 - average ranking)
- Text-based documents (4.03)
- Audio (2.90)
- Moving images (2.67)
- Objects/artifacts (1.29)

3 respondents noted in the comments that they do not prioritize digitization based on content type: priorities are determined by other internal needs, such as subject, collection or client requests.

When asked "**What digital assets do you wish you could preserve but currently cannot?**" (Q62) the highest number of responses related to web-based content (12 responses), including public and internal websites, and websites of student organizations. 2 of these respondents mentioned local digital news sources as a concern. Organizational records, including administrative and operational records from universities and governments ranked second highest (8 responses), followed by research data (5), digital audiovisual carriers (5), e-mail (3), and 3D images (2). 2 respondents each mentioned software, legacy media such as floppy disks, digitized

film, and complex researcher-driven projects with interactive environments or databases.

2 respondents indicated that their current preservation capacities were adequate to preserve the majority of the assets in their collections. In contrast, 2 respondents said that they do not feel they can currently preserve anything in their collections, responding “all of them” to the question, “What digital assets do you wish you could preserve but currently cannot?” 4 others indicated it is not a content-related issue, but an organizational or resource-driven one: 3 of these respondents indicated their main issue is uncertainty around the scope of digital acquisitions and how to assess and prioritize materials for preservation, and 1 mentioned a lack of resources dedicated to the area.

Key takeaways:

- Most respondents are collecting born-digital materials and all are pursuing digitization activities.
- While the content sources and types differ, a higher emphasis on organizational and private records points to the need to support archives-specific workflows for preservation.
- Organizations require support for preserving more complex, multi-faceted and web-based content, as well as research data and e-mail, and materials on digital carriers such as floppy disks.
- Only a small proportion of respondents are confident in their abilities to preserve their holdings, while several are not undertaking any activities to preserve digital assets or do not collect digital materials due to a lack of preservation capacity.

3. Organizational Commitment and Engagement

One indication of an organization’s commitment to digital preservation is the language contained in strategic plans, mission statements, or mandates. **41 respondents (79%) have language that expresses a commitment to digital preservation either published, waiting approval, drafted, or planned** (Q11), while 11 respondents (21%) have no language in place.

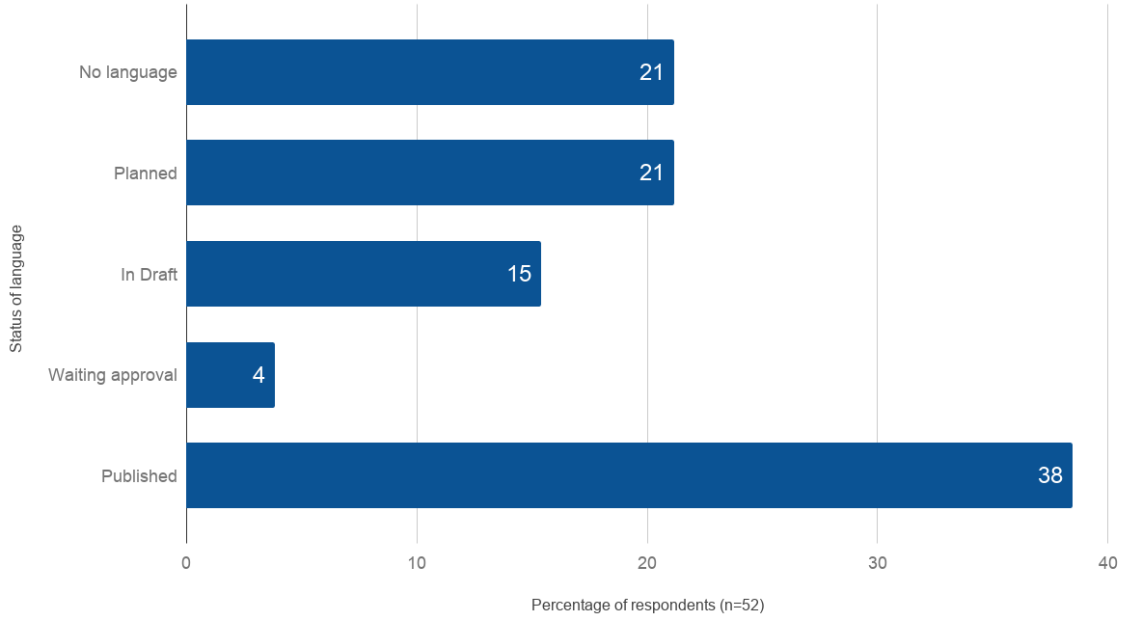

Figure 5: Status of language in a strategic plan or mission/mandate statement expressing commitment to digital preservation ranked by percentage of respondents (Q11).

When asked to list departments or units in which digital preservation activities are taking place (Q12), 28 respondents (54%) indicated that they are taking place in archives and special collections units. 2 respondents (4%) specified records or information management units. 11 (21%) specified information technology or systems units. Academic libraries cited a range of digital scholarship, digital initiatives, and research services units (19 respondents, 37%). 9 respondents cited collections units (17%). 7 respondents (13%) indicated the library itself, and 2 (4%) said "all units." 1 respondent listed a specific unit devoted exclusively to digital preservation. **33 respondents (63%) listed more than one unit or department undertaking digital preservation activities in their organization**.

42 respondents (81%) have an individual or group responsible for coordinating digital preservation activities across the organization (Q13). 10 respondents indicated that there was no group or individual coordinating digital preservation at their organization. Of the 42 positive responses, 12 (29%) indicated that a group was responsible, including 2 senior leadership teams and 3 digital preservation working groups. 16 (38%) indicated that one or two individuals in a senior management role, such as a university librarian, city archivist, dean, associate dean, or the manager or director of a unit, were responsible. 11 (26%) named one or two individuals in an operational role, such as a staff librarian or archivist without an administrative title. Finally, 3 (7%) named a mix of two or more individuals in both senior and operational

roles where it was identified that the senior staff member coordinates activities and the staff member responsible for operations performs the day-to-day work.

Table 1: Examples of titles of individuals or groups responsible for digital preservation (Q13)

Individuals	Working Groups or Committees
• Associate Dean and the Head of University Archives and Special Collections • Associate Dean, Digital Initiatives • Bibliothécaire en chef de la Bibliothèque • Chief Records Officer • City Archivist • Digital Archivist • Digital Collections Administrator • Digital Initiatives Librarian • Digital Preservation Librarian • Director of Access and Conservation Services • Director of Collections Management • Director of Scholarly Technologies • University Archivist • Volunteer Archivist	• Digital Preservation Working Group • Digitization Committee • Digital Strategies • Information Management Steering Committee • Library Technology Services • Senior Management and Leadership Team

23 respondents (44%) have a committee or working group related to digital preservation and 29 respondents do not (Q14). Of the 23 positive responses, 15 (65%) named one or more formalized groups with a focus on digital preservation specifically (Q15). 5 (22%) named formal committees or working groups with interests that may relate to digital preservation, such as digital initiatives, information management, and data governance, but without specific mention of digital preservation. 2 (9%) named an informal group of individuals with interests in the area. 1 respondent (4%) did not name the group. Of the 20 formal groups listed, 18 (90%) had an ongoing mandate (Q16) and 19 (95%) were authorized by a senior individual or leadership team (Q17).

The roles for these groups (Q15) were primarily described as identifying and drafting policies, providing oversight for operations, and ensuring alignment with strategic

plans. Individual respondents identified additional activities such as monitoring trends, identifying best practices for digitization, developing training materials, prioritizing materials for preservation, and inventorying assets.

Many respondent organizations are **participating in external organizations, projects and initiatives related to digital preservation** (Q18). 39 respondents (75%) indicated participation or membership in regional, national, or international organizations, conferences or projects specific to digital preservation. 10 respondents are not participating in such initiatives, while 3 respondents indicated that they did not know about their organization's participation.

The following word cloud illustrates the organizations, conferences, projects and initiatives weighted by the frequency with which they were mentioned by respondents. The highest frequency was 13 times and the lowest was 1 time. 44% of the 49 distinct entities listed were mentioned just once by respondents. See Appendix 3 for a key to acronyms and a complete list of entities listed.

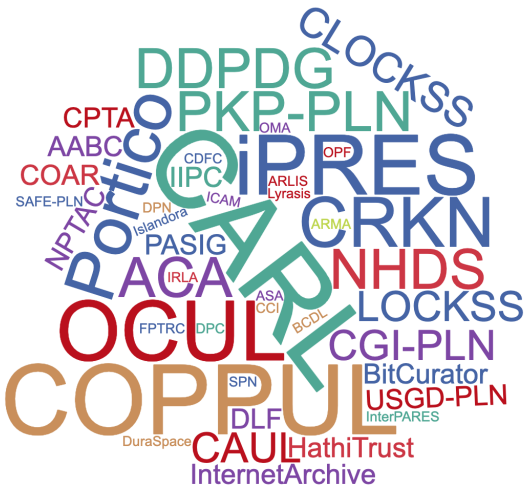

Figure 6: Word cloud of organizations, projects, initiatives or conferences related to digital preservation in which respondents are participants (Q18). The size of the words is determined by the frequency with which they were mentioned by respondents. Please note that this image is not an exhaustive representation of the entities named; see Appendix 3 for a list of all the words and a key to their acronyms and their respective weights.

The **gaps and challenges related to organizational commitment and engagement** (Q19) were:

- Lack of resources to pursue organizational change (33 respondents, 63%)
- Lack of resources to participate in collaborative efforts (26, 50%)

- Lack of communication/coordination among stakeholders (23, 44%)
- Lack of high-level organizational commitment or support (19, 37%)
- No gaps/challenges (3, 6%)

Additional gaps and challenges noted were: a lack of coordination, prioritization or strategic direction at the organization's highest level (8 respondents), awareness of digital preservation is not yet consistent/shared across organization (8), a lack of funding or other resources generally (6), the problem of developing skills or finding digital preservation expertise (3), the high cost of infrastructure and storage (3), the large size, scope or complexity of the materials to be preserved (3), the challenges of moving from ad hoc to more formalized processes (2), and challenges in using shared/centralized services (2). One respondent commented: "Our largest challenge is that our IT infrastructure is done through a centralized IT service for the government, who are ultimately responsible for providing business IT services for all government ministries. As our organization plays a unique role in the preservation and long term access of government records, it is challenging for IT to understand the different role and function of archival records for the purpose of digital preservation versus access and use of business records for the function of running a government agency day-to-day."

Key takeaways:

- Many organizations are in the process of formalizing their role related to digital preservation, though some have already made their commitment public through strategic plans or mission statements.
- Digital preservation activities are taking place in various units and departments, but are often coordinated across the organization by a senior administrator or committee.
- There is a high level of engagement with external organizations, projects and initiatives related to digital preservation reflecting the diverse activity of the field and the need for peer knowledge exchange and resource sharing.
- There is uncertainty about how to shift resources away from other areas towards digital preservation activities, reflecting the challenges organizations may face in presenting business cases to institutional leaders.

4. Policies, Strategies, and Procedures

Policies

All respondents indicated that they have implemented, or are interested in implementing, a digital preservation policy at their organization (Q20). 9 respondents (17%) have an existing policy. 10 more (19%) have policies under review or in draft, and 33 (63%) are either in discussion about policy development or have no policy.

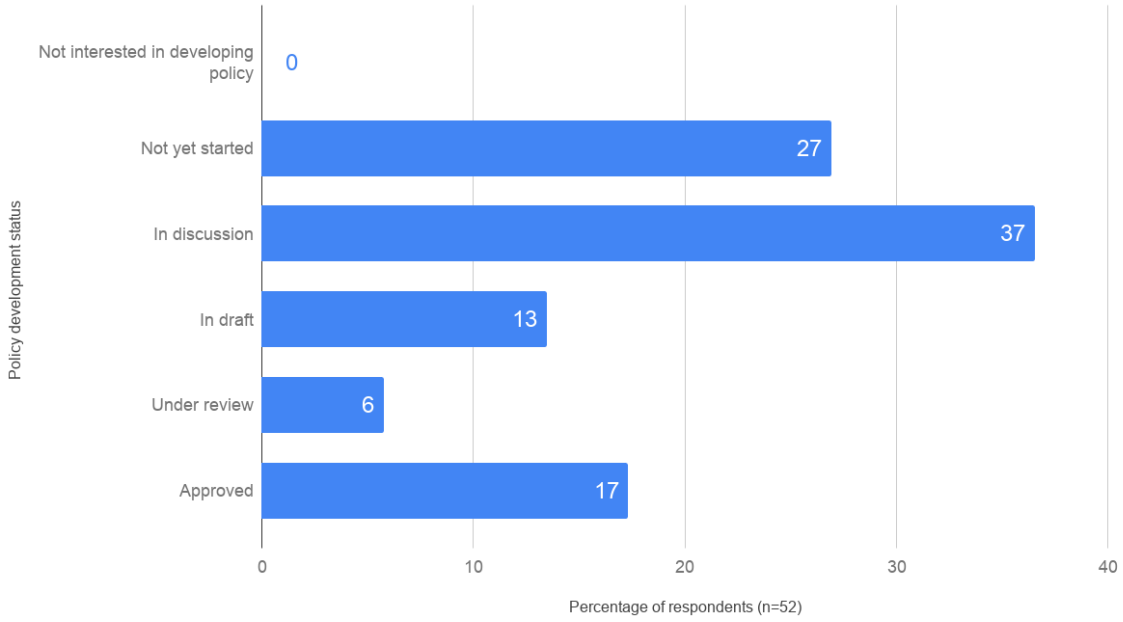

Figure 7: Status of policy development at respondent organizations ranked by percentage of respondents.

Of the 19 respondents who provided details about the scope of their policies (Q20), the following topics were noted:

- 9 policies address the mandate or administrative framework under which digital preservation activities occur, including roles and responsibilities, resourcing, and commitments to long-term preservation.
    - 2 of these address sustainability and cost, and 2 address partnerships with external organizations.
- 5 policies provide collections frameworks such as acquisition, access, and prioritization for preservation.

    - 2 of these address levels of preservation being applied to different content types and 2 address the specific content-related scope of the organization's preservation activities.
- 2 policies include guidelines for transferring digital archival materials, metadata and digitization, storage, security and disaster planning.
- 1 policy consists of a broad framework accompanied by content action plans for different units in the organization.
- 1 policy relates specifically to content in a digital repository.
- 1 policy, according to the respondent, is completed but is outdated.

Table 2: Policy scope examples provided by respondents (Q20)

- This policy applies to all digital preservation activities undertaken within the [organization] system. This policy is intended to outline what can be expected from digital preservation programs and activities at [organization] and to place reasonable limits on those expectations given expected operating constraints and technological considerations. This policy may form the basis for partnership agreements if [organization] provides digital preservation services to non-[organization] partners, now or in the future.

- It defines the legal and administrative framework, the scope and its applicability. It covers both the principles and commitment towards service delivery, and efficient and active preservation of digital records.

- The Digital Preservation Policy applies to all digital objects for which [organization] is the primary custodian, including born-digital and digitized material accessioned into the [organization's] collection or Institutional Archives. This material may arrive at [organization] in any format and on any media. Specific preservation decisions are always made in the larger context of [organization]'s institutional priorities and programs, and are contingent on available resources, the needs of [organizations]'s users, and the perceived research value of materials.

- To identify and prioritize all of the content groups that should be considered for digital preservation. To identify the appropriate level of preservation for each of those groups. To determine and document procedures for each content group. To ensure that selected preservation tools are available, and that the preservation workflows are operationalized.

- [Organization] commits to preserving the materials for which it has accepted responsibility to the greatest degree possible. The primary objective of [organization]'s digital preservation activities is to preserve the intellectual content repository, the representation of digital or digitized records, as well as the content

> metadata. [Organization] will prioritize the preservation of materials ingested into the digital repository as well as legacy digital materials according to the Levels of Preservation documentation.

The **gaps and challenges related to policies** (Q27) were:

- Lack of time/resources for policy development (41 respondents, 79%)
- Policies are ad-hoc or project-specific (21, 40%)
- Lack of knowledge for policy development (12, 23%)
- Policies are not well documented (12, 23%)
- Policies are not reviewed (9, 17%)
- Policies are not well understood or followed (8, 15%)
- Lack of interest in policy development (5, 10%)
- No gaps/challenges (2, 4%)

4 respondents indicated that a major barrier to policy development is a lack of knowledge among staff. One respondent said, "Policy writers in the organization don't understand or lack knowledge regarding digital preservation. Roles in policy writing [are] not well defined in the organization." Another noted, "We plan to develop policies this year. The largest gap seems to be a clear and shared understanding of the scope of what is to be preserved." 4 respondents said that they are awaiting a higher-level organizational shift in order to drive policy development, including the purchase of technical tools or infrastructure that will help clarify policy requirements. 2 respondents noted that, while policies are in place, they are inconsistently applied across units or content types. One respondent also commented: "Policies need to be at a fairly high level with procedures and guidelines applying to many different types of collections and data types managed via different technologies. This can be challenging to manage. Examples include repository and digitized collections vs. Archive-it (web archives) vs. large research datasets."

Strategies or Plans

35 respondents (67%) have digital preservation plans or strategies in place (Q21). 17 respondents said they do not. Of the 35 positive responses, 25 respondents (71%) said that these plans or strategies address file format standards for preservation and access, including content action plans or file format policy registries. Commonly listed formats were text, photographs, audio, and video. In addition, 6 (24%) listed digitization strategies, and 4 (16%) mentioned strategies for storage and backup.

Procedures and Workflows

Many organizations do not yet have consistent procedures and workflows related to digital preservation (Q22), with a majority of respondents indicating their organization has either no procedures, undocumented procedures, or draft procedures (together, 40 respondents, 77%). 12 respondents (23%) indicated they have procedures that are documented and in use.

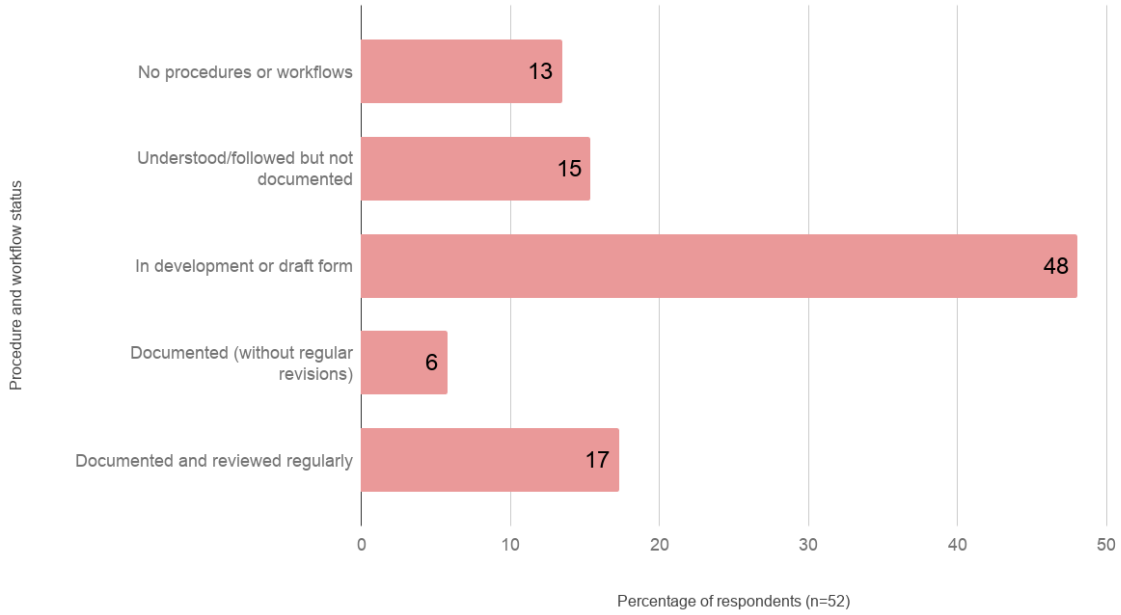

Figure 8: Status of procedures or workflows at respondent organizations ranked by percentage of respondents.

11 respondents listed at least one procedure in use (Q22). These included:

- 6 references to content-specific procedures, including content action plans and format standards.
- 4 references to digitization workflows.
- 3 references to tool-specific sets of workflows, such as Archivematica, Taverna, the use of Bags, and AtoM.
- 3 references to procedures/standards relating to metadata or archival description and processing.

The **gaps and challenges related to procedures and workflows** (Q28) were:

- Lack of time/resources for procedure documentation (39 respondents, 75%)

- Procedures are ad-hoc or project-specific (25, 48%)
- Lack of knowledge for procedure documentation (15, 29%)
- Procedures are not well documented (13, 25%)
- Procedures are not well understood or followed (12, 23%)
- Procedures are not reviewed (10, 19%)
- Lack of interest in procedure documentation (5, 10%)
- No gaps/challenges (3, 6%)

Respondents noted additional challenges related to procedures and workflows:

- 2 respondents said that keeping documentation current is a challenge.
- 2 said that there is a lack of coordination/information-sharing across units.
- 2 indicated a lack of knowledge/skills for procedure development among staff at the organization.
- 2 are awaiting technical infrastructure to direct the structure/content of procedures.

1 respondent commented, "Some audiovisual preservation procedures tend to be fluid as we bring new equipment online and let go [of] older equipment, and due to fluctuating financial and human resources due to changes in government budgeting priorities." Another contextualized their response by noting "This is kind of the 'worst case' scenario. There are plenty of procedures that are well-developed, -documented, and followed, but more are always under development."

Other Policy-related Considerations

30 respondents (58%) indicated that their organization has adopted particular digital preservation standards, best practices or guidelines (Q23). 18 respondents said they have not and 4 respondents said they didn't know. Of the 30 positive responses, the following standards were listed more than once:

- OAIS (13 mentions)
- PREMIS/METS (11)
- TRAC/ISO 16363 (6)
- Dublin Core (4)
- Archivematica's format policies (3)
- MODS (2)
- PAIMAS (2)
- Digitization guidelines:
    - Federal Agencies Digitization Guidelines Initiative (FADGI) (2)

    - Library of Congress guidelines (2)
- The NDSA Levels of Preservation (2)
- Digital Preservation Management Workshop guidelines (2)

21 respondents (40%) indicated they have adopted metadata standards for structuring and managing digital preservation metadata (Q24), 25 said they have not, 4 did not know, and 2 did not respond. The following standards were listed more than once (note that some respondents also listed metadata standards in response to Q23):

- METS/PREMIS (11 mentions)
- Dublin Core (9)
- MODS (3)
- DDI (2)

When asked about **interest in pursuing formal ISO 16363 Trustworthy Digital Repository certification** (Q25), **9 respondents (17%) indicated that they have undertaken or are interested in undertaking formal certification**. None of the respondents have been formally certified to date. 10 respondents said "I don't know" and the remaining 32 respondents answered "No." Of the 32 negative responses, the reasons given were:

- A lack of time/resources (12 mentions)
- Certification is not a priority or seen as providing value (7)
- Interest in meeting the requirements but not obtaining formal certification (6)
- The use of a third party provider who is or will be certified (4)
- The fact that the program in question is not sufficiently developed (4)

1 respondent that indicated "Yes" in the survey was not counted as they qualified their response by stating "Yes, but only through our partnership with selected external services hosted at OCUL's TDR infrastructure. We do not see ISO 16363 as practical at our local level at this time, but are keen to see our larger partners extend compliance." Another respondent noted, "We completed a TRAC self-assessment (audit carried out by Artefactual) and, based on those results, have improved our policies, procedures and documentation and are nearly ready to carry out a second audit. We have no plans to undergo certification by an official body though as the self-assessment is sufficient for our needs. We would be interested in peer review approaches with other CARL Libraries."

15 respondents (29%) have used digital preservation-related self-audit or gap analysis tools (Q26). 30 have not and 7 did not know. The following tools were listed:

- TRAC/ISO 16363 self-audit tools (6 mentions)
- Scholars Portal digital preservation readiness interview (2)
- AVP Cost of Inaction Calculator (1)
- Consultant (1)
- Digital Preservation Management workshop gap analysis (1)
- DRAMBORA (1)
- Internal format monitoring/policy review (1)

Key takeaways:

- While all respondents indicated they are interested in developing policies related to digital preservation, the majority of organizations do not yet have a policy in effect. Most are in the development or discussion stages.
- A larger number of organizations are using formal or draft procedures addressing the practical requirements involved in digital preservation activities, such as processing workflows/guides, handling file formats in repositories and digitization activities.
- The use of standards and assessment tools, which usually indicates a high level of maturity related to digital preservation, is mixed.
- A lack of resources and expertise to develop policies and procedures is a key challenge for many organizations.

5. Tools and Applications for Preservation Functions

Digital Forensics

The survey included a few questions to gauge current usage of digital forensics tools, which enable safer transfer of born-digital materials from external media like hard drives and floppy disks. **29% (15) respondents are employing tools for digital forensics** (Q33). 33 organizations are not using digital forensics tools, and 4 respondents were unsure. The following tools were listed by more than one respondent:

- BitCurator (10 mentions)
- FTK Imager (4)
- Kryoflux (4)
- Tableau write blocker (3)
- FRED (2)
- Guymager (2)

- ISO Buster (2)

See Appendix 5 for a list of all forensics-related tools mentioned.

Respondents were asked, "If applicable, how are you preserving forensic disk images? (Q34):

- Preserve both the disk image and extracted contents (5 responses, 33%)
- Extract contents and delete the disk image (4, 26%)
- Preserve the full disk image (3, 20%)
- Other (3, 20%)

Of the "Other" responses, 2 said a decision has yet to be made, and 1 said that it depends on the content.

The **gaps and challenges related to digital forensics** (Q35) were**:**

- Lack of staff knowledge/skills (28 respondents, 54%)
- Lack of access to software tools (22, 42%)
- Lack of access to hardware (disk drives, write-blockers, etc.) (21, 40%)
- No gaps/challenges (2, 4%)

Other gaps and challenges noted:

- Lack of resources (including staff) (8 mentions)
- Still learning/developing approaches (4)
- Not an organizational priority (2)
- Lack of storage (2)

Preservation Processing Tools

48% (25 respondents) are employing at least one tool for preservation processing, either in testing or in production (Q36). 11 (44%) of these respondents indicated they are using any one of the tools listed in production (Q38). Archivematica is the most common tool in use, with 13 of 25 respondents (52%) using it in testing and 8 of 25 (32%) using it in production.

Table 3: Uses of preservation processing tools, in testing and production and if locally installed or using a hosted service (Q37 and Q38)

Tool	Testing, installed locally	Testing, hosted service	Production, installed locally	Production, hosted service	Testing, installa-tion type not given	Installed locally, status not given	Total
Archive-matica	6	3	4	4	4	0	21
Arkivum	0	0	0	0	0	0	0
Preservica	0	1	0	0	0	0	1
Other*	2	0	3	0	0	2	7

*Other tools listed: BagIt-based tool (1 in testing – installed locally), Islandora (1 in production – installed locally), Locally-developed scripts (2, 1 in production and 1 not specified - both installed locally), Taverna workflow engine (1 in production - installed locally), The Museum System (TMS) (1 installed locally, status not specified), Ultima from CGI (1 in testing – installed locally).

The **gaps and challenges related to preservation processing tools** (Q39) were:

- Lack of money to support tools (28 respondents, 54%)
- Lack of software/tool support (27, 52%)
- Lack of staff knowledge/skills (27, 52%)
- Lack of access to hardware (16, 31%)
- No gaps/challenges (1, 2%)

Other gaps and challenges noted were:

- Lack of time/staff to pursue this work (9 mentions)
- Organization is still in planning/development stage (8)
- Not currently an organizational priority (3)

One government-based museum noted: “Museums, as expected, prioritize tools to help manage large physical collections. As such, resources towards digital collection tools is generally a secondary priority. Although this is changing as we collect and create more digital objects, the infrastructure and tools to manage digital collections

are not yet on par with those used to [manage] the physical collections." Another CARL member summarized the issues at play: "The conversation regarding building resources and capacity for digital preservation began with becoming a primary partner in the OLRC and now having access to storage space via this partnership. However, [organization] hasn't been actively using this storage as of yet. The intention is to use it in a planned way so that we are actively preserving what needs to be preserved rather than just uploading content without good oversight. As such, this oversight involves clarifying long-term management strategies and policies/workflows. There is also interest in being more effective around access to make the user experience more seamless. Our challenges include devoting staff resources, obtaining sustainable long-term storage and infrastructure, formalization of policies, and development of partnerships locally, regionally, and nationally."

Key takeaways:

- A minority of organizations are using digital forensics processes and digital preservation processing tools as part of their routine work.
- Preservation processing-related tools for such functions as file format identification, characterization, and normalization enable more comprehensive approaches in digital preservation workflows. The low use of tools indicates that the majority of respondents are not performing preservation processing functions for the digital materials in their care.
- The low adoption of tools is largely due to a lack of funding and staffing.

6. Discovery and Access

96% of the respondents (50) indicated that they provide access to digital materials under their care (Q29). 2 organizations (both community-based/non-profit organizations) do not give access to digital materials. Respondent organizations are providing access via the following methods (respondents could select more than one option):

- Web platform/repository (44 respondents, 85%)
- Web transfer/shared folder (27, 52%)
- Dedicated computer terminal (21, 40%)
- We do not give access/not applicable (2, 4%)
- Other (12, 23%)

Respondents who said "other" are using the following methods for providing access:

- Files on external media (CDs, USBs, etc.) (3 mentions)
- Website (general) (3)

- E-mail (2)
- Direct contact with staff person (2)
- Social media (1)
- Centralized online resources (HathiTrust, Internet Archive, Canadiana, etc.) (1)
- Purchase requests (1)
- FOI requests (1)

79% of respondent organizations (41) use a digital object repository or discovery platform (Q30) and 11 do not. Respondents were asked to identify the type of repository system/software in use and how it is used:

Table 4: Access systems in use (Q31)

System	Digital objects only*	Descriptions only	Digital Objects and Descriptions	Not specified
ArchivesSpace		1		
AtoM		9	16	
Blacklight			1	
ContentDM			7	
Dataverse	4	1	14	
DSpace	1		17	
Islandora	1		11	
Samvera/Hydra			3	
Other**	1	4	13	3
Total	7	15	82	3

*May be interpreted to mean uncurated metadata and/or the use of storage repository functions only.
**See Appendix 4 for a list of other access repositories named

Note that there was a gap of 4 respondents who indicated that they use a web platform/repository in Q29 but did not specify a particular repository or discovery platform in Q30.

The **gaps and challenges related to access** (Q32) were:

- Lack of technological infrastructure (27 respondents, 52%)

- System/software limitations (27, 52%)
- Lack of policies/procedures (22, 42%)
- Lack of storage space (19, 37%)
- Privacy/security issues (19, 37%)
- No gaps/challenges (5, 10%)

Additional challenges noted were:

- Rights/permissions issues (6 mentions)
- Storage costs or difficulties in estimating and expanding storage (6)
- Lack of resources and staff (4)
- Challenges in platform migrations (4)
- Access workflows/approaches still in development (2)
- Specific content type needs not being met for audiovisual materials (2), traditional knowledge (1), and big data (1)

Key takeaways:

- A strong majority of respondent organizations give access to digital materials using web-based platforms. However, a large proportion of organizations also use web transfer methods like shared folders or give access onsite.
- A variety of access platforms are in use.
- A majority of respondents indicated that a lack of access to infrastructure and system/software limitations are barriers to improving access.

7. Storage

Methods

Reliable storage is a key component of a resilient digital preservation program. Respondent organizations are using a variety of storage options, though **most organizations are using local network storage as one of their options (94%)** (Q40). Fewer organizations are using networked storage infrastructures that often provide more reliability, such as cloud networks and tape backups.

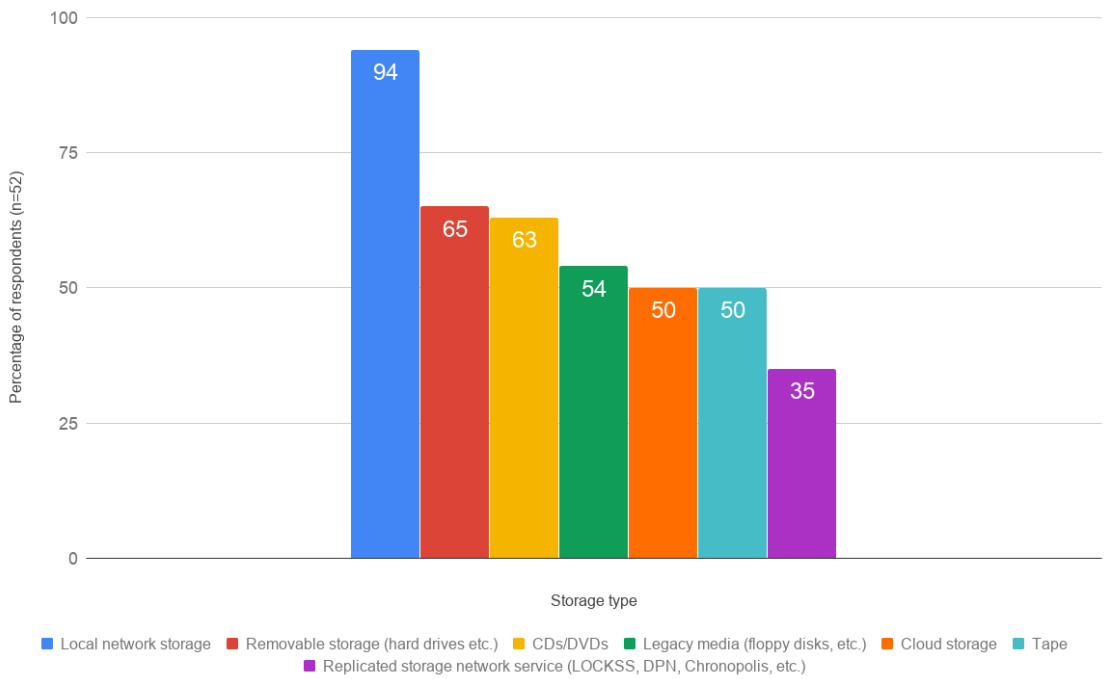

Figure 9: Storage methods selected by percentage of respondents (Q40). Respondents could select more than one option.

All but two of the respondents are using multiple storage methods and media. **The average number of storage methods used was 4.15**. Responses show a high reliance on storage media such as removable hard drives, CDs and DVDs, and legacy media. 81% selected at least one of these options, and 38% selected all three. In contrast, the storage methods more widely considered appropriate for preservation (cloud, tape, and replicated storage services) fall at 50% usage or below.

Of the 26 respondents (50%) using cloud storage the following types of cloud storage providers were indicated (Q41):

- Private/community cloud (9 responses, 35%)
- Both private/community cloud and commercial provider (8, 31%)
- Commercial cloud provider (6, 23%)
- Both private/community cloud and Intermediary/hybrid (e.g. DuraCloud) (1, 4%)
- Not indicated (2, 8%)

Providers that were named by more than one respondent were:

- Ontario Library Research Cloud (8 mentions)
- Amazon (5)
- Microsoft (2)
- Google (2)
- UBC EduCloud (2)

The use of cloud storage is almost exclusive to academic libraries, with only 2 (8%) other respondent organizations using cloud storage (a community/non-profit organization and a government-based archives).

Of the 18 respondents (35%) using replicated storage networks like LOCKSS (Q42), the following methods/providers were selected or listed by respondents:

- LOCKSS (14)
- CLOCKSS (5)
- WestVault (2)
- Digital Preservation Network (2)*
- PKP-PLN (1)
- SAFE-PLN (1)
- APTrust (0)
- Chronopolis (0)
- Not indicated (1)

*Both respondents are members of COPPUL, so this may be interpreted to mean COPPUL's Digital Stewardship Network.

Note: 5 respondents listed the OLRC in response to Q42 and were excluded. Two respondents did not select "Replicated storage networks" in Q40 but added responses in Q42 and these entries were included in the chart for Q40.

Similar to cloud storage, the use of replicated storage methods is largely limited to academic libraries. Of the 18 respondents who selected replicated storage networks as an option, a single government-based archives represented an organization outside of the academic sector.

Quantities

65% of respondents (34) were able to quantify the amount of digitized content their organization stores (Q58). The median amount was 16 TB and average was 29.72 TB with a low of 500 GB and a high of 200 TB. Storage quantities fell into the following three broad distributions:

- 0-10 TB (15 respondents)
- 10-100 TB (15)
- 100-200 TB (4)

48% of respondents (25) were able to quantify the amount of born-digital content their organization stores (Q59). The median amount was 4.4 TB and average was 28.59 TB with a low of 50 GB and a high of 330 TB. Storage quantities fell into the following distributions:

- 0-10 TB (14 respondents)
- 10-60 TB (9)
- 130 TB (1)
- 330 TB (1)

In Q60, Respondents were asked to estimate the **total amount of content stored** if they answered "unknown" for Q58 or Q59. 21 responses were given, 7 of which still could not indicate a number (13% of respondents overall). Of the 45 respondents who were able to give two numbers in Q58 or Q59, or a total number in Q60, **the median amount of total data stored was 20 TB and the average was 133.29 TB**. Total storage quantities fell into the following distributions:

- 0-10 TB (17 respondents)
- 10-100 TB (21)
- 100-400 TB (6)
- 3,720 TB (1)

The following represents total average amounts of storage divided by sector:

- Academic-based respondents outside of CARL: 29.5 TB, 2 unknown
- CARL members: 62.40 TB, 2 unknown
- Community/non-profits: 7.15 TB, no unknowns
- Government-based organizations: 425.45 TB, 3 unknowns
    - When the large outlier of 3,720 TB is removed, the average for government respondents is 59.4 TB.

Note on the data in this section: if respondents entered a range (e.g. "10-20 TB") the upper end of the range was counted as the figure.

Distributions for Born-Digital Content

65% (34) of respondents were able to estimate the distribution of born-digital data across storage methods (Q61). Those who responded indicated on average that 65% of assets were stored on networked systems and 33% on external media. External media could include external hard drives and disks from donors, as well as the use of similar media for storage by the organization itself. A smaller proportion of 12 respondents indicated they had 2% of assets, on average, stored on internal media, such as a donor's personal computer.

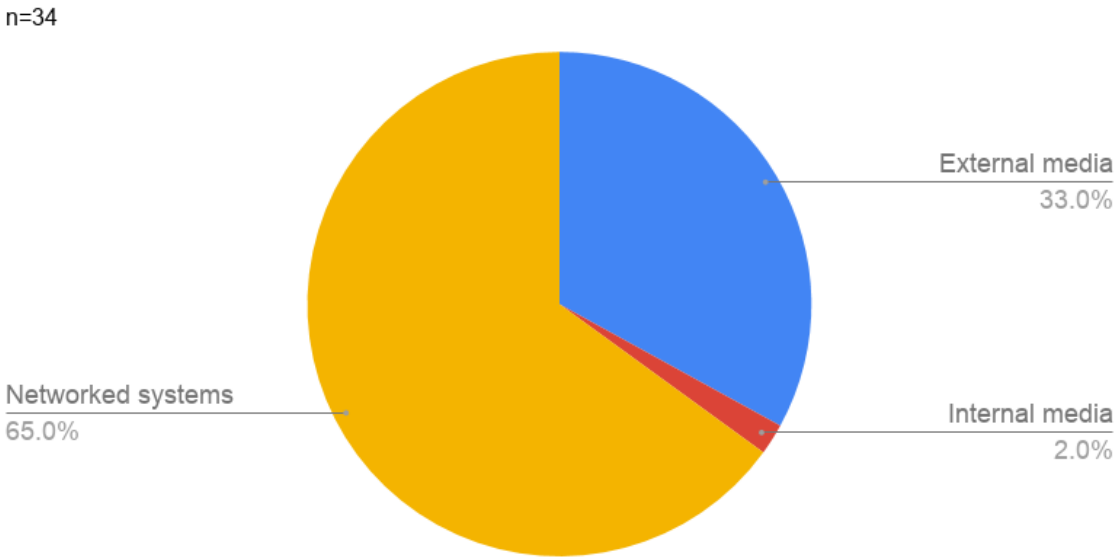

Figure 10: Distributions of stored content based on broad media type (Q61).

Note on the data in this section: if the percentages did not add to 100, the response was disregarded.

The **gaps and challenges related to storage** (Q43) were:

- High cost of local storage (26 respondents, 50%)
- Lack of local storage (18, 35%)
- Lack of oversight/control over storage (15, 29%)
- Lack of support for storage (12, 23%)
- Security/privacy barriers (12, 23%)
- Procurement barriers (11, 21%)
- Lack of backups (10, 19%)
- No gaps/challenges (9, 17%)

In the comments, 5 respondents noted it can be difficult to work with IT units or form collaborative partnerships regarding storage within organizations. One respondent commented, "Our campus ITS is stretched to fulfill their typical responsibilities. We cannot count on them or local resources to provide digital asset storage at our campus. We are too small for this to scale properly." 4 respondents also noted the interrelated issues of cost and scalability/flexibility, including difficulties in estimating costs.

Key takeaways:

- Most respondents are using local network storage to keep at least some of their assets. There remains widespread reliance on external media and legacy media for storage. Reliance on these media present significant risks where media are not backed up, can be prone to hardware error, may degrade quickly over time, or are potentially already obsolete.
- Resources for the transfer of data from legacy materials and hard drives, including forensics capabilities, will be required to ensure the safe transfer to more reliable storage systems.
- The transition to preservation-friendly storage among all respondents has been slow. A better understanding of the barriers to adoption in this area is needed.
- Respondents had difficulty estimating amounts of data stored based on type, though most could give an approximate total amount. One respondent commented on the difficulty of distinguishing between storage for born-digital and digitized content: "I would expect that few libraries that have extensive digital collections can easily determine how much disk storage each type of content consumes. Answering this question would require more sophisticated digital assets management capabilities than most libraries are capable of." This points to a need for improved tools for the management and oversight of digital assets across storage systems and content types.
- High local storage costs remain an issue for half of the respondents.

8. Staffing

Number of Roles

96% of respondents (50) have 1 or more roles with some responsibilities for digital preservation (Q44), with the majority of organizations (74%) having 3 or more roles with some level of responsibility.

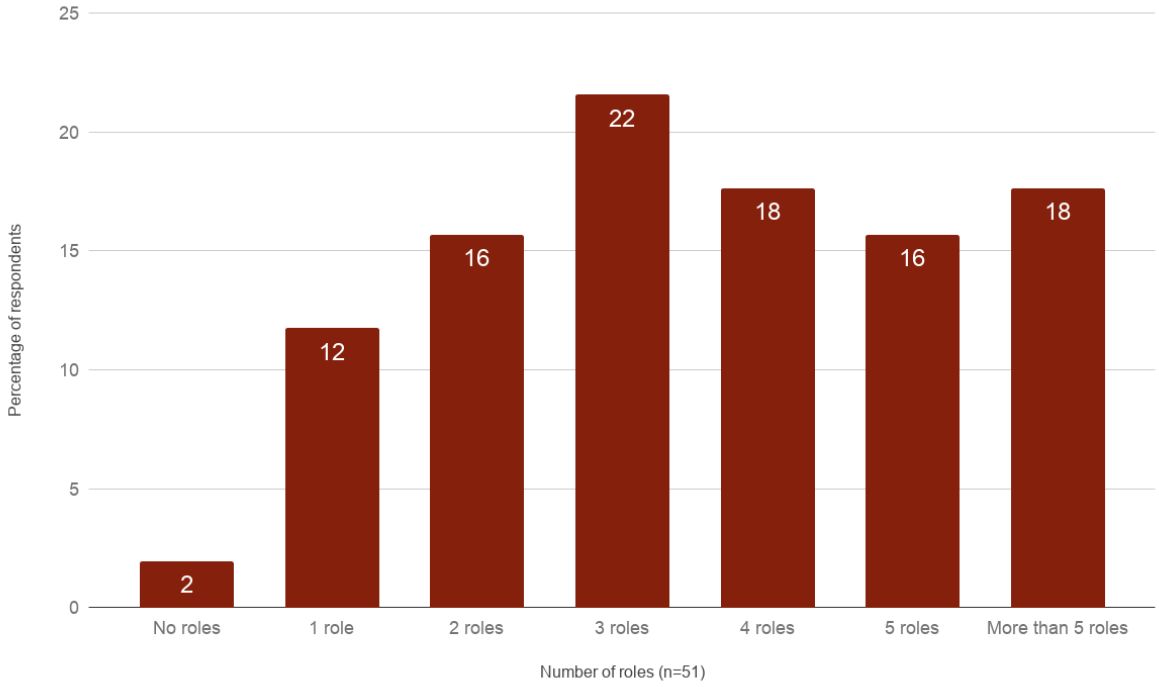

Figure 11: Number of roles dedicated to digital preservation at respondent organizations ranked by percentage (Q44). Note: 1 respondent indicated "unknown" for this question and was not counted in the percentage values.

Job titles for digital preservation-related roles varied widely (Q44). Of the 170 job titles entered by respondents, the following titles formed clusters:

- Archivists: 18
- Digital Archivists (including "Digital Curation Archivist," Digital Project Archivist"): 10
- Variations on "Digital Librarian" (e.g. "Digital Projects Librarian," "Digital Repository Librarian," and "Digital Initiatives Librarian"): 19
    - Digital Preservation Librarian: 3 of the above
- Systems Librarians: 4
- IT Staff (Developers, Systems Administrators, Programmers): 13
- Senior staff*: 47

*University Librarians and/or Archivists, City Archivists, Associate University Librarians/Deans, and heads of units or other management staff ("senior," "manager," "coordinator," or "director" in their title).

FTE Values

FTE values provided by respondents shows a **low level of staff time directed toward digital preservation work** (Q44):

- Of the 49 respondents who listed FTE values for one or more roles, 40 (82%) do not have at least one full-time role dedicated to digital preservation.
- Across all roles listed, 31 (62%) do not have staff roles with digital preservation responsibilities that add up to 100% FTE.

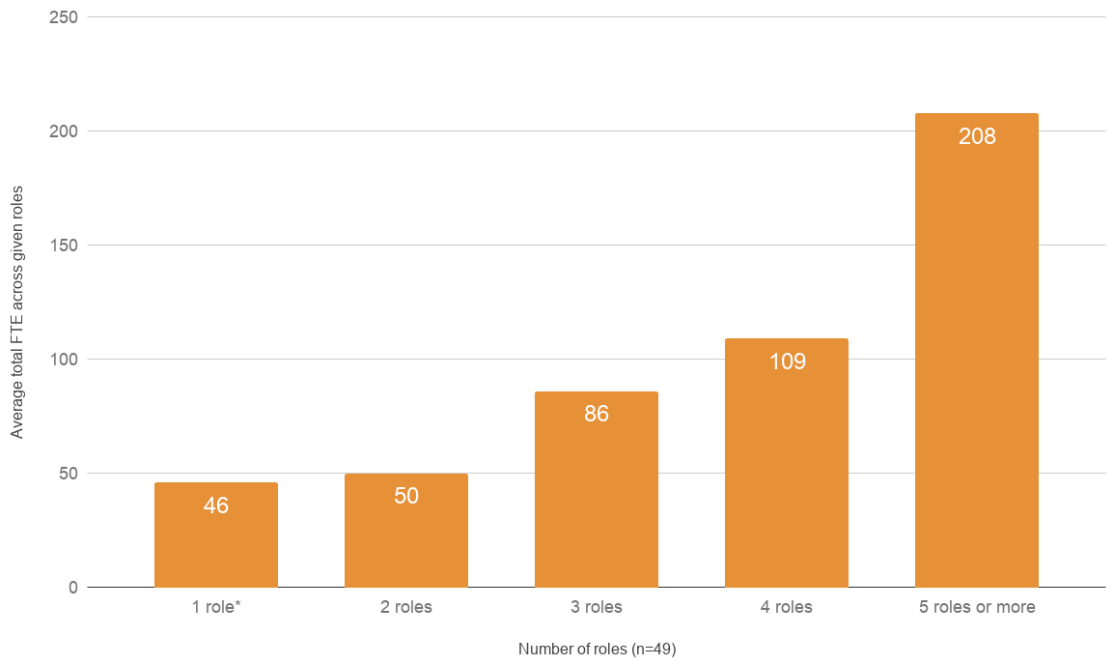

Figure 12: Average total percent FTE based on number of roles at a respondent organization (Q44). *Note: also counted as "one role" in this chart were respondents that listed more than one job title but only one FTE value. Note that two respondents were disregarded as they did not include FTE values and one respondent with no dedicated staff was not counted.

For the 34 respondents with 1-4 roles listed, the average total FTE across these roles was 0.77 FTE per respondent. This average rises to 1.17 FTE when counting the 17 respondents with 5 roles or more, including the 4 respondents with 5 full-time roles (n=49).

Of the 166 listed roles with FTE values assigned:

- 48% (80 roles) had between 0 and 20% FTE assigned
- 36% (60 roles) were between 20 and 100% FTE

- 16% (26 roles) were 100% FTE roles.

Overall, the average FTE for digital preservation per respondent (including the respondent with no roles, n=50) is 1.14 FTE.

One additional conclusion that can be drawn from the data is the relationship between FTEs for digital preservation and overall FTEs at an organization. The 27 CARL member respondents that participated in the survey reported a total library FTE of 4,792.89 in 2016-17.[1] The total digital preservation-related FTE for these same respondents in the survey was 29.85, or 0.62% of FTE overall.

The 22 Phase 2 respondents who both indicated a total FTE value for their organization in Q7, as well as a value for digital preservation staffing, reported a total of 2,505.12 FTE. The total digital preservation-related FTE was 26.74, or 1.06% of FTE. In total, this means that of the 7,298.01 total organizational FTEs reported across 49 respondents, **FTEs for digital preservation represent 0.77% of all positions.**

Expectations for extending staff responsibilities for digital preservation were mixed (Q45). 25 respondents (48%) said they intended to expand staffing. 12 said "No" and 15 said "I don't know."*

The following methods for increasing staffing were selected (Q46):

- Both new hires and assigning current staff (10 respondents, 40%)
- Reassigning current staff only (8, 32%)
- New hires only (7, 28%)

*If a respondent selected "I don't know" in Q45 this meant that any additional selections they may have made were not counted in the statistics for Q46.

The **gaps and challenges related to staffing** (Q47) were:

- Lack of funding for new positions (40 respondents, 77%)
- Lack of staff knowledge/skills (23, 44%)
- Lack of resources for training/professional development (17, 33%)
- No gaps/challenges (1, 2%)

[1] Canadian Association of Research Libraries. (2018). *CARL Statistics, 2016-17*. [link].

A few respondents listed additional gaps and challenges. 2 respondents noted that since their programs are in development, it is hard to identify needs and advocate for staff. One respondent commented: "There are currently no FTE positions remaining for the organization, therefore any additional staffing will need to be filled through contract positions (budget allowing) as we have done with most other functions for the department." Another said, "It is very long and complex to acquire the necessary skills to gain basic knowledge of digital preservation and no training is readily available. This means a new hire often has to be trained from scratch and the field of digital preservation is vast!"

Key takeaways:

- Most respondents have low staffing levels devoted to digital preservation. While many organizations have at least one or two individuals with some responsibilities in the area, the average is the equivalent of 1.14 full time employees, though actual responsibilities are likely to be divided into small portions among a number of employees. 62% of organizations have less than 1 FTE responsible for digital preservation in total across all staff.
- Many organizations intend to expand staffing through new hires and reassignment of existing staff.
- The major challenge is a lack of funding for positions.
- Additional challenges are related to expertise, whether in terms of training existing staff or finding new staff with the necessary skills and knowledge.

9. Funding

85% of respondents (44) are funding digital preservation through general budgets (Q48). Respondents selected the following funding sources (respondents could select more than one):

- General budgets (44 responses, 85%)
- Grants or awards (22, 42%)
- IT budget (16, 31%)
- Gifts or endowments (13, 25%)
- Materials budgets (10, 19%)
- Preservation budget line (7, 13%)
- Other (3, 6%)

"Other" responses included a collections budget, computing budget, and an archives store.

3 respondents listed grants or awards as the sole source of funding for digital preservation.

75% of respondents (39) could not indicate what percentage of their organization's budget was dedicated to digital preservation (Q49). The 13 remaining responses grouped as follows:

- Of the 10 respondents that gave percentages:
    - Less than 1% (4)
    - 0.3% (1)
    - 3-3.5% (2)
    - Less than 5% (1)
    - 8-8.6% (2)
- Of the 2 respondents that gave dollar figures: one respondent each listed $5,000, $15,000, $98,000 (including student salaries)

32 respondents (62%) indicated they expected budget increases in the next 1-2 years beyond normal increases (Q50). 8 respondents (15%) said they did not expect increases, and 12 (23%) said they didn't know. However, 78% (25) of those who responded "yes" to the question could not indicate what percentage value this expected increase would be. Of the 7 responses with a value assigned, respondents said they expected increases of:

- 10-20% (4)
- 100% (1)
- 200% (1)
- "100K for new hire" (1)

Indication of **resources coming from other stakeholders** was mixed (Q51). 11 respondents (21%) indicated "none."

Table 5: Sources of external resources for digital preservation activities (Q51). Respondents could select more than one option and add additional options.

Internal units	External organizations
IT department (34, 65%)	Regional consortium (19, 37%)
Legal department (6, 12%)	National consortium (10, 19%)
Records management/archives units at universities (2, 4%)	Professional association (7, 13%)
Broader government ministry (1, 2%)	Regional or peer stakeholders (2, 4%)

The **gaps and challenges related to funding** (Q52) were:

- Allocation of resources is too low in comparison to needs (27 respondents, 52%)
- Funding is not sustained (26, 50%)
- Lack of business plan (21, 40%)
- No gaps/challenges (3, 6%)

4 respondents noted in the comments that programs are in development, and funding will hopefully follow clearer determination of needs or a firmer commitment to digital preservation on the part of the organization. One respondent commented that "Many projects are driven by one-time funding, but commitments are long-term." Another noted "The 'back office' association with the various activities can sometimes be challenging in contrast to other activities with more public visibility."

Key takeaways:

- While general budgets are a key source of funding, the relatively high reliance on short-term funding such as grants and awards for a long-term activity like digital preservation is cause for concern.
- Most respondents had difficulty in estimating how much money is directed to digital preservation work.
- Increases in funding over the next 1-2 years are expected by the majority of respondent organizations, though the extent of these increases is not known.
- The majority of organizations rely on resources from IT units, but fewer respondents noted resources from external organizations.
- Most respondents from all sectors feel that funding resources are inadequate.
- There are additional funding challenges related to difficulties in estimating costs and advocating for needs based on unknown costs.

10. Organizational Maturity

Respondents were asked to **rank the status of their digital preservation programs in general terms** using a standard maturity scale,[2] from 0 (No activity) through to 5 (Optimized - processes are mature and continually improved) (Q63). 72% of respondents placed themselves at stages 1 or 2, 27% at 3 or 4, and none at stage 5.

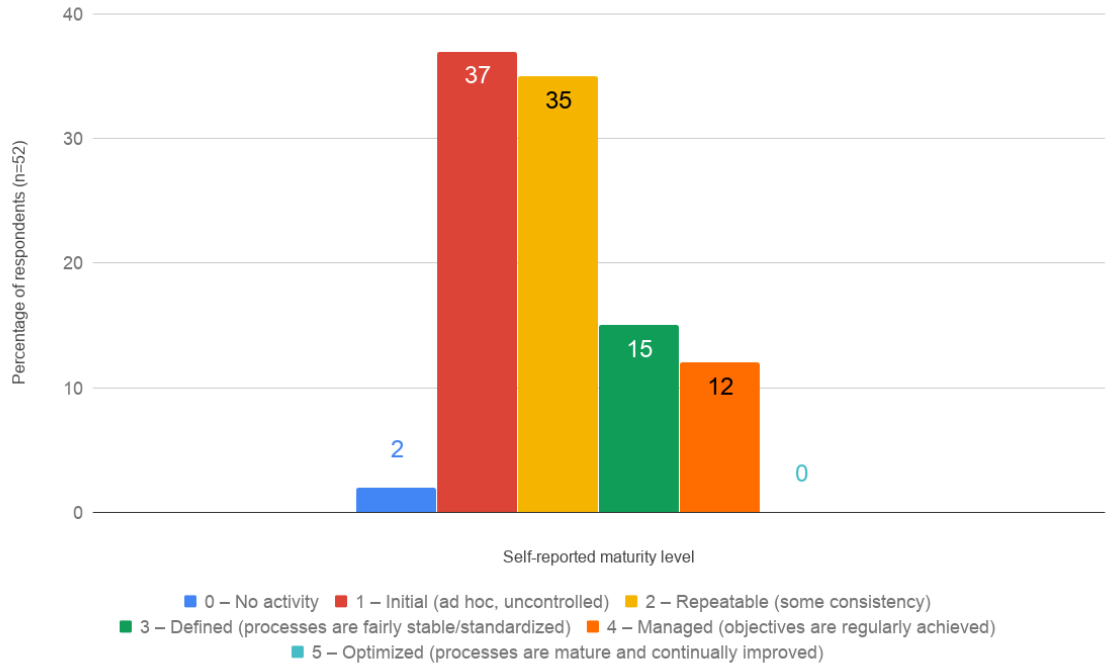

Figure 13: Self-reported maturity scales ranked by percentage of respondents.

Respondents finished the survey with the opportunity to add additional comments in response to the question "**Are other organizational, policy, technological, and resource issues preventing digital preservation capacity?**" (Q64). Aside from expressing a general lack of resources and funding (16 respondents), 8 respondents noted that policy issues are a barrier; 6 indicated staff knowledge and time; 5 indicated competing priorities at the organization, 4 mentioned storage barriers (lack of access and cost), and 3 mentioned challenges to coordination and collaboration.

[2] Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). Capability maturity model, version 1.1. *IEEE software*, *10*(4), 18-27. [link].

Table 6: Selected comments in response to "Are other organizational, policy, technological, and resource issues preventing digital preservation capacity?" (Q64)

- The budgetary constraints of the last years impact our capacity to ensure the integrity and prevent the development of our digital collection. Resources were primarily affected. Without adequate funding to ensure long-term preservation of our documentary heritage, [deleted] it is our entire collective memory that is at risk.
- Our biggest challenge is the costs associated and fear of these costs, staff time to concentrate on preservation activities, and Canadian only consortium to participate in due to government data storage regulations which prevent our participation with more established partnership groups.
- Finding knowledgeable staff is a challenge continuously.
- We need to go slow on consortial efforts to allow for enough time for the collective effort to coalesce.
- Lack of a dedicated policy office, lack of a dedicated individual for intellectual property, lack of a digital preservation policy, lack of leadership in these areas at a high level, lack of resources needed for digitization.
- 1. Centralized IT service provider does not understand digital preservation requirements and cannot provide solutions needed.

  2. Organizationally, other parts of organization can override digital preservation decisions and recommendations.

  3. Funding fluctuates depending on current management in place.
- We have the technology and an organizational will to increase our capacity for digital preservation. Formulating a policy or creating a group or roles with specific responsibilities to do this work will be needed to expand our digital preservation activities.
- No money and low interest in increasing staff time to do digital preservation work.
- Most significantly, we need to create the time to develop policy and procedures. Resources (storage, software) are not significant barriers at this time.

Key takeaways:

- A strong majority of respondents (72%) do not yet consider their digital preservation activities to be mature, while a smaller group (27%) feel they are mature as measured by the adoption of formalized processes and objectives.

E. Analysis: Relative Capacity

1. Scoring the Survey Results

As an exploratory measure, respondents were scored across key areas of the survey that contained quantitative information: organizational commitment; policies and procedures; tools and technologies for preservation and access; storage; and staffing and funding. Appendix 2 contains the scoring rubric. The intention was not to assess respondents against a preexisting standard, but to identify areas of strengths and weaknesses and provide for a higher-level set of comparisons in the results.

The method used was to assign points that represented levels of investment or progress in each area. Each of the five sections, weighted equally, was allotted up to 3 points to make for a maximum total of 15 points. The scores reflect relative capacity across the responses and therefore should not be taken as absolute values. For example, a score of 14 or 15 out of 15 points indicates high capacity within the cohort of respondents, but not some absolute potential capacity for that organization. Moreover, this survey does not measure the portion of content being "preserved," as appropriate preservation approaches will vary based on the defined needs of the organization, the types of content to be preserved, and the resources at hand. Because different content types require different strategies and resources, and preservation approaches are always in flux, digital preservation programs are best evaluated based on the sets of functions and activities they are able to support relative to their specific needs.

Of the organizations surveyed, the lowest score was 3.5 and the highest was 15; the average score overall was 9.17. The average scores per sector were similar:

- Academic respondents outside of CARL: 9.43
- CARL members: 9.72
- Community/non-profit: 8
- Government: 8.5

The following chart shows the distribution of average scores.

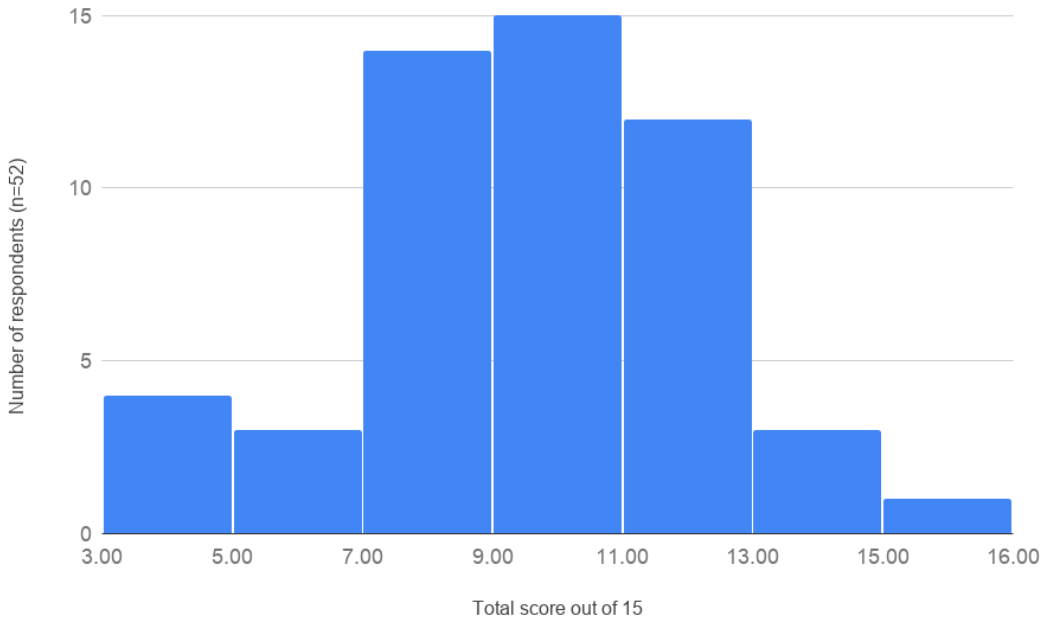

Figure 14: Distribution of scores out of 15 by number of respondents.

Table 7: Score distribution against relative capacity out of 100.

Score range	Number of respondents	% of respondents	% relative capacity
0-3	0	0%	0-19%
3-5	4	8%	20-33%
5-7	3	6%	33-46%
7-9	14	27%	47-59%
9-11	15	29%	60-73%
11-13	12	23%	74-86%
13-15	3	6%	87-99%
15	1	2%	100%

Areas of relative strength and weakness across the five areas:

- Organization & Governance (average of 2.30 out of 3)
- Storage (2.17)
- Policies & Procedures (1.95)
- Access/Tools for Preservation (1.55)
- Staffing and Funding (1.19)

The large proportion of respondents fell into the middle cohort of scores, with smaller groups of low and high scoring respondents on either side. The general score cohorts are divided as follows:

- A cohort of 7 respondents (13%) in their initial stages, have a relative capacity of between 0-46%. 3 of these respondents had strengths in relation to storage, but otherwise had low or no scores in other areas. This group consisted of 3 CARL academic libraries, 2 community/non-profits, and 2 government-based public libraries.
- A large middle group of 41 respondents (79%) with programs in development had a relative capacity from 47-86%. This group contained the majority of respondents, and their areas of strength and weakness reflects the averages provided above. This group consisted of 21 CARL libraries, 11 government-based organizations, all 7 of the non-CARL academic libraries, and 2 community/non-profit groups.
- A small group of 4 respondents (8%) had high relative capacity from 87-100%. This group consisted of 3 CARL libraries and 1 community/non-profit group.

When mapped to the maturity areas in the "Concluding Comments" section, there is a general observable correlation with the "initial" and "repeatable" areas of maturity for the respondents in the middle cohort. The higher-capacity cohort consisted of 2 respondents that rated themselves at the Defined level (#3) and another 2 at the Managed level (#4).

2. Score Comparisons Against Organizational Profile Data

Another potential outstanding question about the results is the relationship between total organizational resources and capacity for digital preservation. The following two sections briefly present some comparisons.

CARL members

The following charts were prepared using data from CARL's 2016-17 statistical report.[3] They show the relative capacity scores against two values: materials expenditures and total library staff FTE. The comparisons show that there is not a definite relationship between higher materials expenditures and staff FTEs and capacity in digital preservation. While 3 organizations with relatively high materials expenditures and staffing levels come out on top in terms of their scores, the other top 10 organizations do not necessarily reflect the highest collections budgets or FTE staff numbers. Of the top 10 scorers (with scores between 11.5 and 15), 5 fall below the median materials expenditure value for all respondents of $10,772,038, and 4 fall below the median total FTE value of 145.3. However, on the other end, the bottom 3 scorers have expenditures and total FTE levels either around or below the median levels. Therefore, while a large, well-resourced organization can help support a strong digital preservation program, it is not a prerequisite.

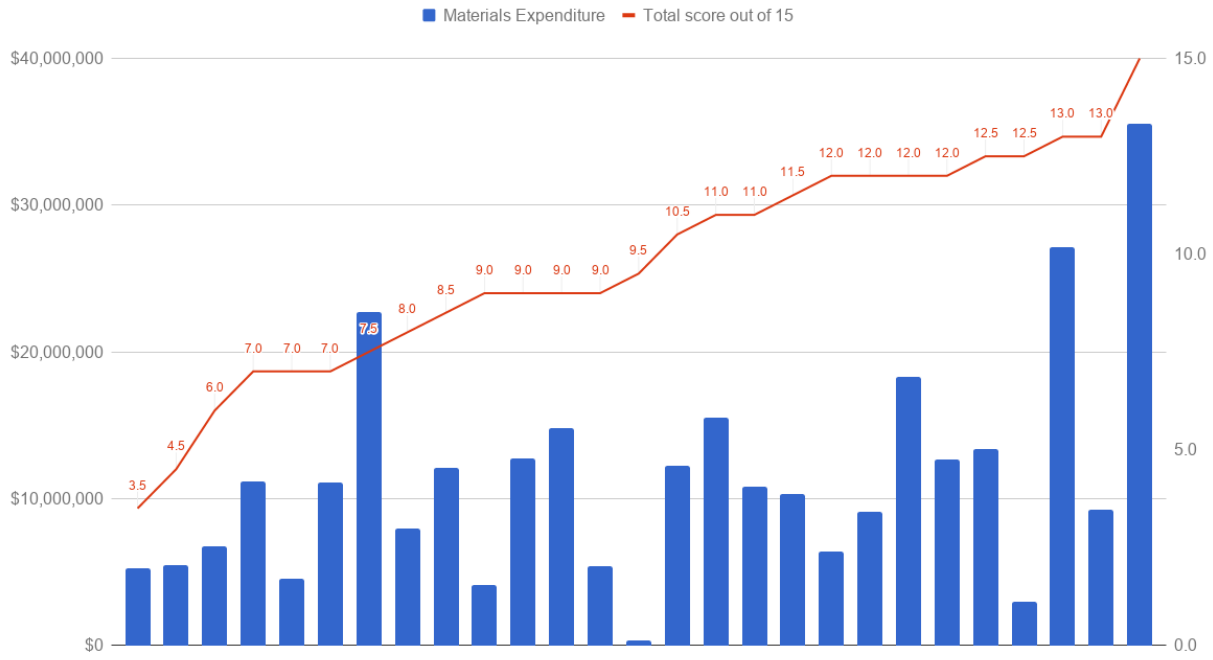

Figure 15: CARL member materials expenditures vs. relative capacity score (n=27).

[3] Canadian Association of Research Libraries. (2018). *CARL Statistics, 2016-17*. [link].

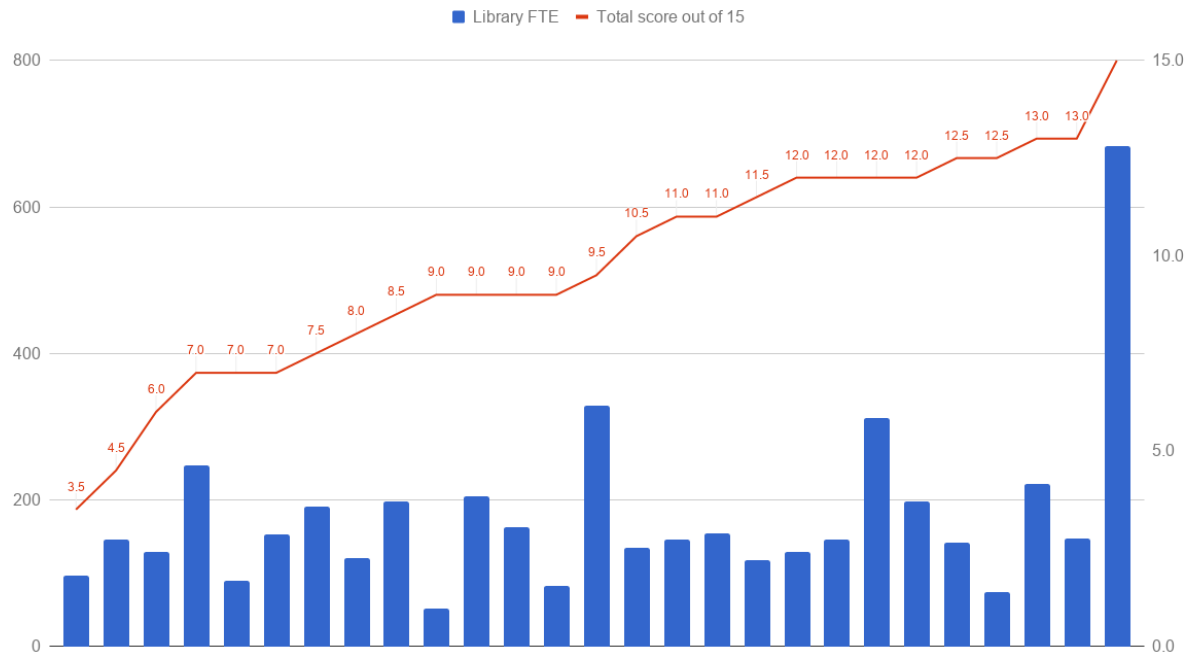

Figure 16: CARL member FTE values vs. relative capacity score (n=27).

Phase 2 Respondents

Comparison charts for Phase 2 respondents show a similar set of results: organizations with the largest budgets or staffing values do not necessarily have the highest scores related to digital preservation.

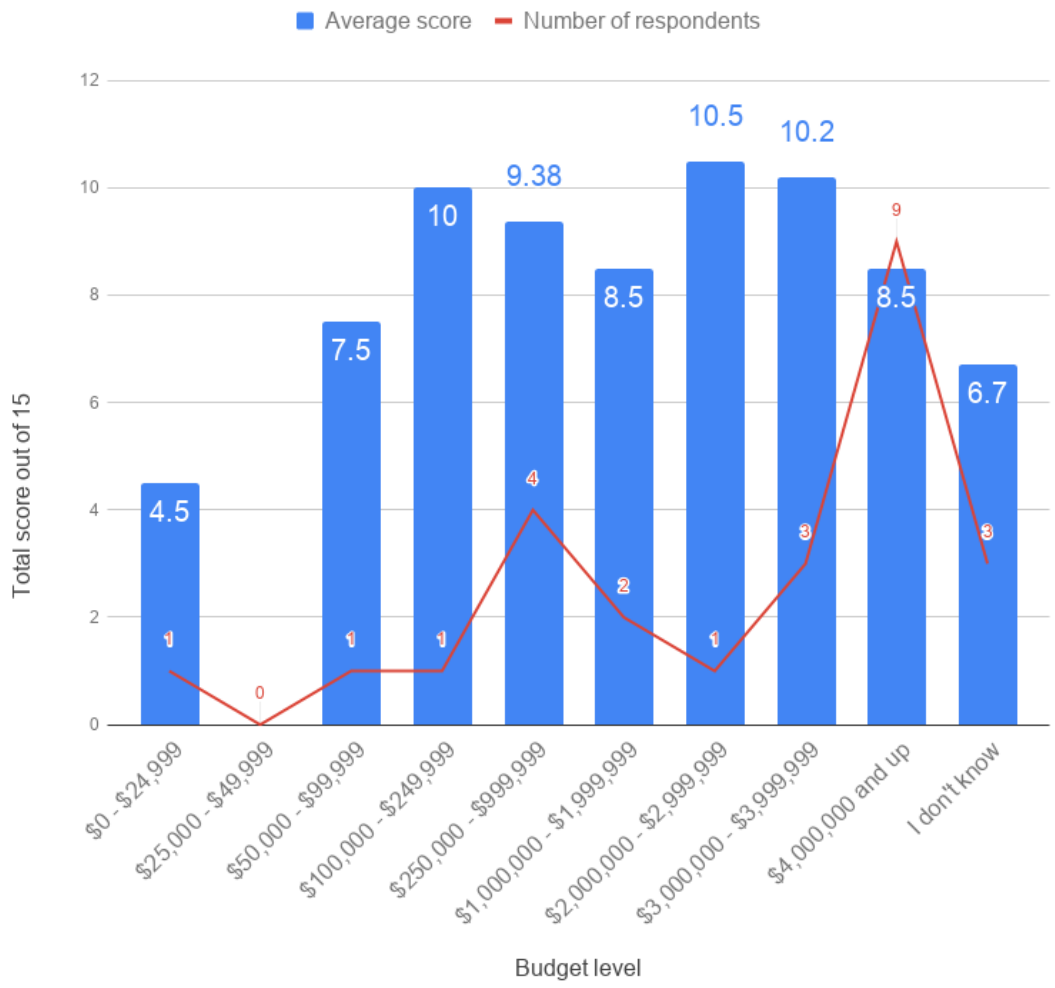

Figure 17: Budget range against average relative capacity score out of 15 for Phase 2 respondents. The red line indicates the number of respondents per budget category (n=25).

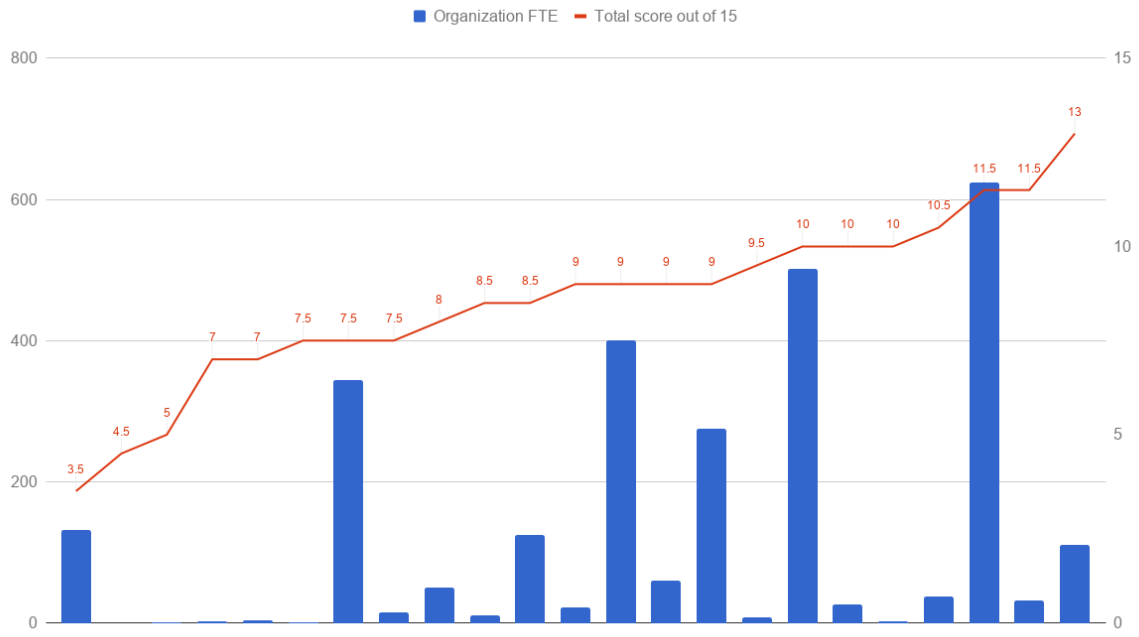

Figure 18: Self-reported total organization FTE values expressed by Phase 2 respondents against relative capacity score. Two respondents said "I don't know" in this question (n=23).

F. Conclusion

The aim of this survey was to understand the current state of digital preservation in Canada in the library, archive, gallery and museum sectors. The results show that there are still many challenges for Canada's memory institutions related to digital preservation. This is a real concern, as the volumes of digital heritage, education and research content are growing with each year. Despite missions and mandates to ensure preservation and access over the long-term, many memory organizations have not yet prioritized digital preservation within their operations and do not direct significant resources towards these activities. Digital preservation expertise is also a critical deficiency for many. Because digital preservation is complex and multifaceted, it requires a range of skills and knowledge spread across a variety of positions, from senior administrators who need to set policy directions, to systems positions who are responsible for storage and backup, to metadata specialists and technicians. Moreover, as technology changes, so do the methods and processes related to digital preservation, making it difficult for staff to remain up-to-date with current practices.

The survey demonstrates that many institutions, both large and small and across the different sectors, are struggling with similar challenges related to digital preservation,

and collective action in Canada could be very helpful in addressing some of the issues. Shared storage, technologies, and staffing, for example, could potentially fill in a number of gaps. In recognition of these challenges, there are already existing initiatives pursuing collaborative approaches, such as those being undertaken by the regional library associations, the Portage Network, and others. However, some institutions are not covered through these efforts and there is much more that could be done, especially in collaboration with organizations outside of the academic library community.

In order to address the numerous staffing challenges, training and skills development covering a range of aspects related to digital preservation (e.g. policy, organizational, and technologies) that are accessible to all memory institutions in Canada would be beneficial. Communities of practice that may evolve out of training activities could present an opportunity for ongoing knowledge sharing, enabling institutions to continually learn from others and contribute their experiences to the collective knowledge.

Memory organizations, which are known more for their enduring nature rather than their agility and ability to adapt quickly over time, could also benefit from strategic support to help manage their transition from analogue to a mixed analogue/digital ecosystem. The high-level case for digital preservation still needs to be articulated to decision makers that may have little understanding of the requirements, costs and issues on the ground. Organizational models will need to change to reflect new priorities related to digital preservation. Developing best practice and modern organizational models in the digital age that more appropriately reflect digital preservation needs can provide the impetus for institutions to begin the restructuring process that is required.

Canada's digital cultural, historical, research and educational content is an invaluable resource for our country and it is critical that we collectively ensure that this content is available to Canadians now and in the future. To address the many challenges, there is a key strategic, leadership role for national and regional organizations to help organizations increase their local capacity related to digital preservation, as well as to act as mechanisms for shared services and infrastructure. Additionally, given the scope of the challenges and the numerous players in the ecosystem, there is also a need for coordination across organizations. We suggest that LAC and CARL consider how to strengthen the national effort by defining and staffing a national digital preservation coordination role, perhaps under the aegis of an existing organization. This role would help focus the collective activities and their impact for libraries and archives, but could extend as necessary across the entire GLAMs sector.

G. Acknowledgements

First and foremost, the authors would like to thank the respondents who took time out of their day-to-day work to gather information for the survey. There would be little to report without your contributions! Many thanks go to Lise Brin and Julie Morin at CARL for their tireless assistance and support in administering the survey. The authors also thank the members of the DPWG Readiness Survey Subgroup for their assistance and editorial advice in the design of the survey and the writing of this report: Lise Brin, Alan Darnell, Corey Davis, Susan Haigh, Steve Marks, and Michael Moosberger. We also thank the broader DPWG membership for their feedback on this report, and for their support of the survey initiative overall. All remaining errors in this document are the authors' own.

H. Appendices

Appendix 1: Survey Questions

Building Capacity for Digital Preservation in Canada: Survey on Current State and Future Needs. A survey by Canadian Association of Research Libraries (CARL)'s Digital Preservation Working Group – Readiness Subgroup

To preview this questionnaire as a PDF file, please click here.

Introduction

The digital preservation landscape is complex and rapidly evolving, and memory institutions are increasingly looking for solutions to address the substantial challenges associated with their emerging digital stewardship responsibilities. The CARL Digital Preservation Working Group (DPWG) seeks input from libraries and archives across Canada on the state of digital preservation capacity and gaps in their organizations. [Added in Phase 2] In 2017, the DPWG surveyed CARL member institutions and organizations reflected in the DPWG; this second phase targets archives, libraries, galleries, and museums outside of CARL's membership.

The goal of the survey is to provide for a more accurate and updated picture of digital preservation activities in Canada, and will help to identify existing gaps and outstanding needs at Canadian institutions. Ultimately, it is anticipated that the findings will contribute to the development of strategies, policies, expertise and resource allocation that will enable the community to build capacity and help to ensure that Canada's valuable digital assets are preserved for future generations. The aggregate survey results will be made available via a white paper report. All survey data will be anonymized.

Definitions

Organization: the particular group on whose behalf you are answering this survey. The unit of inquiry could include an entire organization or a unit within a larger organization. [Added in Phase 2] For example, you could respond as an archives unit within a municipality, or on behalf of the municipality as a whole. However, please note that the wider the scope of the organization, the more information your responses will need to contain.

Digital assets: materials held by an organization whose primary form is digital and whose value for evidential, cultural, social, historical, or other purposes is considered eligible for long-term management for access and preservation.

Digital preservation is defined as "the series of managed activities necessary to ensure continued access to digital materials for as long as necessary" (Digital Preservation Coalition, 2016).

[Added in Phase 2] Please ensure that you respond on behalf of your organization, rather than as an individual. This is to ensure that appropriate data is collected in compliance with section 2.5 of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans.

This survey will take 30-60 minutes to complete. You may save your entries and return to the survey at any time.

If you have any questions, please contact Lise Brin at lise.brin@carl-abrc.ca.

A. Introductory Information

Note: Your "organization" is the particular group on whose behalf you are answering this survey. The unit of inquiry could include an entire organization or a unit within a larger organization. [Added in Phase 2] For example, you could respond as an archives unit within a municipality, or on behalf of the municipality as a whole. However, please note that the wider the scope of the organization, the more information your responses will need to contain.

Q1. [Mandatory] Please provide your name:
(Note: Data will be anonymized. This information will be used to ensure that there is no duplication of information from each institution.)

Q2. [Mandatory] Please provide your email address:
(We may contact you for clarification on your responses, and will let you know once the results are released.)

Q3. [Mandatory] What is your title?

Q4. [Mandatory] Select organization type: [Options: Academic institution; Corporate/Private; Government: Municipal, Provincial/Territorial, National; Community organization/Non-profit; Other (please specify):]

Q4.1 [Mandatory if 'Other' selected] Other (please specify):

Q5. [Mandatory] Select organization subtype: (Select more than one response for combination organizations) [Options: Library; Archives; Gallery; Museum; Other (please specify):]

Q5.1 [Mandatory if 'Other' selected] Other (please specify):

Q6. [Mandatory] Please select the appropriate range for your organization's overall annual budget: [Options: $0 - $24,990; $25,000 - $49,999; $50,000 - $99,999; $100,000 - $249,999; $250,000 - $999,999; $1,000,000 - $1,999,999; $2,000,000 - $2,999,999; $3,000,000 - $3,999,999; $4,000,000 and up; Other:; I don't know]

Q6.1 [Mandatory if 'Other' selected] Other:

Q7. [Phase 2 only] If applicable, please indicate your organization's total yearly materials expenditure.

Q8. [Mandatory, Phase 2 only] Please indicate your organization's total number of personnel expressed as Full Time Equivalent (FTE). (Indicate "unknown" if you are unable to ascertain this figure.)

Q9. [Mandatory] Organization name:

Q10. [Mandatory] What is your organization's general collection mandate/focus?

B. Organization & Governance

Q11. [Mandatory] What is the current state of your organization's commitment to digital preservation, as expressed through language or wording in a strategic plan or mission statement? [Options: No language in strategic plan or mission statement; Adding language is planned; Updated language is being drafted; Language is in place, but awaiting approval; Language is published and available; Other (please indicate)]

Q11.1 [Mandatory, if 'Other' selected] Other (please indicate)

Q12. [Mandatory] In what units or areas, to the best of your knowledge, are digital preservation activities occurring across your organization?

Q13. [Mandatory] Is an individual or group responsible for coordinating these activities? [Options: Yes; No; I don't know]

Q13.1 [Mandatory if 'Yes' selected]: If yes, what is their role?

Q14. [Mandatory] Does your organization have a committee or working group responsible for digital preservation or an aspect of digital preservation activities (e.g. policy)? [Options: Yes; No; I don't know]

Q15. If applicable, what is the committee or working group's title and their role/goals?

Q16. If applicable, is the committee or working group's mandate ongoing or limited? [Options: Ongoing; Limited; Other (please specify):]

Q16.1 [Mandatory if 'Other' selected] Other (please specify):

Q17. If applicable, what individual or unit is responsible for authorizing/initiating the group?

Q18. [Mandatory] Is your organization a member of, or participant in, any regional, national, or international organizations, conferences or projects specific to digital preservation? [Options: Yes; No; I don't know]

Q18.1 [Mandatory if 'Yes' selected]: Please list them.

Q19. What gaps or challenges exist for your organization regarding support for digital preservation at the organizational and governance level? (Select all that apply, and add additional items). [Options: Lack of high-level organizational commitment or support; Lack of communication/coordination among stakeholders; Lack of resources to pursue organizational change; Lack of resources to participate in collaborative efforts; No gaps/challenges]

Q19.1 Please list any additional gaps or challenges, or comments on the above:

C. Policies & Procedures

Q20. [Mandatory] What is the status of your organization's digital preservation policy? [Options: Not interested in developing policy; Not yet started; In discussion; In draft; Under review; Approved]

Q20.1 [Mandatory if 'In draft,' 'Under review' or 'Approved' selected]: What is the scope of the policy? (For example, what high-level activities does it address?)

Q21. [Mandatory] Are any digital preservation plans and strategies in place? [Options: Yes; No]

Q21.1 [Mandatory if 'Yes' selected]: What content types or formats do they address?

Q22. [Mandatory] What is the status of digital preservation procedures or workflows? [Options: No procedures or workflows; Understood/followed but not documented; In development or draft form; Documented (without regular revisions); Documented and reviewed regularly]

Q22.1 [Mandatory if 'Documented (without regular revisions)' or 'Documented and reviewed regularly' selected]: If procedures are documented, please list them, or comment on any of the above.

Q23. [Mandatory] Has the institution adopted any particular digital preservation standards, best practices or guidelines? [Options: Yes; No; I don't know]

Q23.1 [Mandatory if 'Yes' selected]: Please list these.

Q24. [Mandatory] Has the institution adopted any metadata standards for structuring and managing digital preservation metadata? [Options: Yes, No, I don't know]

Q24.1 [Mandatory if 'Yes' selected]: Please list these.

Q25. [Mandatory] Has the organization undertaken, or is interested in undertaking, ISO 16363 Trustworthy Digital Repository Certification? [Options: Yes; No; I don't know]

Q25.1 [Mandatory if 'No' selected]: Why not?

Q26. [Mandatory] Has the organization used any digital preservation-related self-audit or gap analysis tools? [Options: Yes; No; I don't know]

Q26.1 [Mandatory if 'Yes' selected]: Please list them.

Q27. What gaps or challenges exist regarding the status of digital preservation policies? (Select all that apply.) [Options: Lack of interest in policy development; Lack of time/resources for policy development; Lack of knowledge for policy development; Policies are ad-hoc or project specific; Policies are not well understood or followed; Policies are not well documented; Policies are not reviewed; No gaps/challenges]

Q27.1 Please list any additional gaps or challenges, or comments on the above:

Q28. What gaps or challenges exist regarding the status of digital preservation procedures? (Select all that apply.) [Options: Lack of interest in procedure documentation; Lack of time/resources for procedure documentation; Lack of knowledge for procedure documentation; Procedures are ad-hoc or project specific; Procedures are not well understood or followed; Procedures are not well documented; Procedures are not reviewed; No gaps/challenges]

Q28.1 Please list any additional gaps or challenges, or comments on the above:

D. Technology and Infrastructure - Access

Q29. [Mandatory] How does your organization give access to digital materials? (Select all that apply). [Options: Dedicated computer terminal; Web platform/repository; Web transfer/shared folder; Other (Please indicate); We do not give access/not applicable]

Q29.1 [Mandatory if 'Other' selected] Other (Please indicate)

Q30. [Mandatory] Does your organization use any digital object repository/discovery systems (e.g. ArchivesSpace, AtoM, Blacklight, ContentDM, Dataverse, DSpace)? [Options: Yes; No]

Q31. [Matrix] If applicable, for each digital object repository system in use at your organization, indicate whether it holds digital objects, descriptions, or both. [Options: ArchivesSpace; ArchivesSpace; AtoM; Blacklight; ContentDM; Dataverse; DSpace; Islandora; Samvera/Hydra; Other (Please indicate) as “digital objects” and/or “descriptions”]

Q31.1 [Mandatory if ‘Other’ selected]: List any additional digital object repository systems in use at your organization, and indicate whether these hold digital objects, metadata/descriptions, or both.

Q32. What gaps or challenges exist regarding providing access to digital assets? (Select all that apply.) [Options: Lack of policies/procedures; Lack of technological infrastructure; Lack of storage space; Privacy/security issues; System/software limitations; No gaps/challenges]

Q32.1 Please list any additional gaps or challenges, or comments on the above:

E. Technology and Infrastructure - Tools

Q33. [Mandatory] Is your organization creating forensic disk images? [Options: Yes; No; I don’t know]

Q33.1 [Mandatory if ‘Yes’ selected]: Please list the forensics tools used.

Q34. If applicable, how are you preserving forensic disk images? [Options: Preserve the full disk image; Extract contents and delete the disk image; Preserve both the disk image and extracted contents; Other method (please indicate)]

Q34.1 [Mandatory if 'Other' selected] Please indicate:

Q35. What gaps or challenges exist regarding digital forensics capabilities at your organization? (Select all that apply.) [Options: Lack of staff knowledge/skills; Lack of access to software tools; Lack of access to hardware (disk drives, write-blockers, etc.); No gaps/challenges]

Q35.1 Please list any additional gaps or challenges, or comments on the above:

Q36. [Mandatory] Does your organization use digital preservation-related processing tools (e.g. Archivematica, Arkivum, Preservica)? [Options: Yes; No; I don’t know]

Q37. [Matrix] If applicable, for each digital preservation-related processing tool that your organization uses, please indicate whether the tool is installed locally or accessed via a hosted service. [Options: Archivematica; Arkivum; Preservica; Other as “installed locally” or “hosted service”]

Q37.1 Please list any additional digital preservation-related processing tools in use at your organization, and indicate whether these are installed locally or are a hosted service.

Q38. [Matrix] If applicable, for each digital preservation-related processing tool that your organization uses, please indicate whether the tool is in testing or in production. [Options: Archivematica; Arkivum; Preservica; Other as “in testing” or “in production”]

Q38.1 Please list any additional digital preservation-related processing tools in use at your organization, and indicate whether they are in testing or in production.

Q39. What gaps or challenges exist regarding digital preservation tool capabilities at your organization? (Select all that apply.) [Options: Lack of staff knowledge/skills; Lack of software/tool support; Lack of access to hardware; Lack of money to support tools; No gaps/challenges]

Q39.1 Please list any additional gaps or challenges, or comments on the above:

F. Technology and Infrastructure - Storage

Q40. [Mandatory] In what storage systems/media are digital assets currently kept? (Select all that apply.) [Options: CDs/DVDs; Cloud storage; Local network storage; Legacy media (floppy disks, etc.); Replicated storage network service (LOCKSS, DPN, Chronopolis, etc.); Removable storage (hard drives etc.); Tape; Other]

Q40.1 [Mandatory if 'Other' selected] Please specify:

Q41. If applicable, what type of cloud storage is in use: [Options: Private/community cloud; Commercial cloud provider; Intermediary/hybrid (e.g. DuraCloud)]

Q41.1 [If any of the above selected]: What entity(entities) is(are) the provider(s) of cloud storage?

Q42. If applicable, which replicated storage network service(s) is/are used: [Options: AP Trust; Chronopolis; CLOCKSS; Digital Preservation Network; LOCKSS; Other (please specify):]

Q42.1 [Mandatory if 'Other' selected] Please specify:

Q43. What gaps or challenges exist regarding digital asset storage capabilities at your organization? (Select all that apply.) [Options: Lack of local storage; High cost of local storage; Lack of support for storage; Lack of oversight/control over storage; Procurement barriers; Security/privacy barriers; Lack of backups; No gaps/challenges]

Q43.1 Please list any additional gaps or challenges, or comments on the above:

G. Staff & Resources

Q44. [At least one response mandatory] List all roles that are responsible/accountable for day-to-day digital preservation activities in your organization and indicate what percentage FTE of each is devoted to digital preservation. Please give FTE in whole percentage values, e.g. 50, 100. (Please estimate to the best of your ability.) [Options: for 1-5 roles, list role name and %FTE].

Q44.1 If more than 5, please list additional roles below:

Q45. [Mandatory] Is your organization intending to expand staff responsibilities for digital preservation? [Options: Yes; No; I don't know]

Q46. If your organization is intending to expand staff responsibilities for digital preservation, how will these responsibilities be staffed? (Please check all that apply.) [Options: New hire(s); Reassigning current staff; Other]

Q46.1 [Mandatory if 'Other' selected] Please indicate:

Q47. What gaps or challenges exist regarding staffing for digital preservation responsibilities at your organization? (Select all that apply.) [Options: Lack of funding for new positions; Lack

of staff knowledge/skills; Lack of resources for training/professional development; No gaps/challenges]

Q47.1 Please list any additional gaps or challenges, or comments on the above:

Q48. [Mandatory] What are the funding sources for digital preservation activities at your organization? (Select all that apply). [Options: General budget; Grants or awards; IT budget; Gifts or endowments; Preservation budget line; Materials budgets; Other; I don't know]

Q48.1 [Mandatory if 'Other' selected] Please indicate:

Q49. [Mandatory] Approximately what percentage of your organization's last completed fiscal year expenditures were dedicated to digital preservation (including salaries, storage costs, tools and technologies, etc.)? (Indicate "unknown" if you are unable to ascertain this figure.)

Q50. [Mandatory] Do you anticipate an increase in expenditures dedicated to digital preservation in the next 1-2 years, beyond any normal annual budgetary increases? [Options: Yes; No; I don't know]

Q50.1 [If 'Yes' selected]: By what anticipated percentage? (Indicate "unknown" if you are unable to ascertain this figure.)

Q51. [Mandatory] What other stakeholders/partners contribute resources to digital preservation at your organization? (Select all that apply.) [Options: IT department; Legal department; National consortium; Professional association; Regional consortium; None; Other]

Q51.1 [Mandatory if 'Other' selected] Please indicate:

Q52. What gaps or challenges exist regarding funding for digital preservation activities at your organization? (Select all that apply.) [Options: Funding is not sustained; Lack of business plan; Allocation of resources is too low in comparison to needs; No gaps/challenges]

Q52.1 Please list any additional gaps or challenges, or comments on the above:

H. Digital Holdings & Activities

Q53. [Mandatory] Is your organization collecting/accepting born-digital materials? [Options: Yes; No]

Q54. If your organization is collecting/accepting born-digital materials, from where are they sourced? (Select all that apply) [Options: Government (publicly accessible materials); Faculty members/researchers (e.g. monographs, articles, research data and/or other research outputs); Harvested web content; Your organization (e.g. administrative and operational records, organizational publications); Private donors; Purchases from commercial providers (e.g. journals, books, datasets, etc.); Students; Other (Please indicate)]

Q54.1 [Mandatory if 'Other' selected] Other (Please indicate)

Q55. If your organization is collecting born-digital content, what are their content types? (Select all that apply) [Options: Administrative/operational records; Artworks; Audio; Databases; Government publications and data; Purchased datasets (geospatial, statistical, etc.); Purchased publications (books, journals, newspapers); Open access publications (books, journals); Moving images; Research data; Still images; Theses; Web archives or web-harvested

materials; Private papers; Other (Please indicate)]

Q55.1 [Mandatory if 'Other' selected] Other (Please indicate)

Q56. [Mandatory] Is your organization digitizing materials? [Options: Yes; No]

Q57. [Matrix] If you are digitizing materials, rank your organization's priorities based on the following content types from highest to lowest. [Options: rank Audio; Moving images; Objects/Artifacts (3D scanning); Photographs; Text-based documents as "1 (lowest priority)," "2," "3," "4," or "5 (highest priority)"]

Q58. [Mandatory] Approximately how many terabytes of digitized content has your organization created? (Indicate "unknown" if you are unable to ascertain this figure.)

Q59. [Mandatory] Approximately how many terabytes of born-digital content has your organization collected? (Indicate "unknown" if you are unable to ascertain this figure.)

Q60. If "unknown" was indicated for one or both of the questions above, please indicate the total approximate amount of digital content your organization holds in terabytes. (Indicate "unknown" if you are unable to ascertain this figure.)

Q61. Approximately what percentage of born-digital content is stored on each of the following types of storage: (Indicate "unknown" if you are unable to ascertain this figure.) [Options: External media (magnetic, optical, hard drives, flash); Internal media (i.e. donor's personal computer); Networked systems]

Q62. [Mandatory] What digital assets do you wish you could preserve, but are currently not being preserved?

I. Concluding Questions

Q63. [Mandatory] On a scale of 0-5, please rate how your organization is doing overall in terms of preserving digital assets. [Options: 0 - No activity; 1 - Initial (ad hoc, uncontrolled); 2 - Repeatable (some consistency); 3 - Defined (processes are fairly stable/standardized); 4 - Managed (objectives are regularly achieved); 5 - Optimized (processes are mature and continually improved)]

Q63.1 Please add any elaborating comments:

Q64. [Mandatory] Are other organizational, policy, technological, and resource issues preventing digital preservation capacity?

Q65. Do you wish to receive a copy of your responses? [Options: Yes; No]

Q65.1 [If 'Yes' selected]: Please supply the email address where this will be sent:

Q66. Please add any final comments.

Sources consulted

Algana, L. et al. (2017). Beyond the repository: Survey on distributed preservation. [link].

CHIN. (2011). Digital preservation survey: 2011 Preliminary results. [link].

Hurley, G. (2016). Permafrost: Digital preservation readiness and needs assessment interview guide. [link].

McMillan, G., Schultz, M. & Skinner, K. (2011). SPEC Kit 325: Digital preservation. [link].

McMillan, G., Schultz, M. & Skinner, K. (2012). PLN members and the ARL digital preservation survey preliminary findings [link].

Meyer, A. (2015). An annual survey towards a collaborative digital preservation strategy for Canada. [link]

ULCC. (n.d.). Assessing organisational readiness toolkit.

Appendix 2: Relative Capacity Comparison Scoring Rubric

B. Organization & Governance

Q11. What is the current state of your organization's commitment to digital preservation, as expressed through language or wording in a strategic plan or mission statement?

No language in strategic plan or mission statement - 0 points
Adding language is planned - 0.5
Updated language is being drafted - 1
Language is in place, but awaiting approval - 1
Language is published and available - 1

Q13. Is an individual or group responsible for coordinating these activities? ***or*** *Q14. Does your organization have a committee or working group responsible for digital preservation or an aspect of digital preservation activities (e.g. policy)?*

"Yes" in either column counts as 1 point; "No" or "I don't know" as 0.

Q18. Is your organization a member of, or participant in, any regional, national, or international organizations, conferences or projects specific to digital preservation?

"Yes" counts as 1 point; "No" or "I don't know" as 0.

C. Policies & Procedures

Q20. What is the status of your organization's digital preservation policy?

Not interested in developing policy - 0 points
Not yet started - 0
In discussion - 0.5
In draft - 1
Under review - 1
Approved - 1

Q22. What is the status of digital preservation procedures or workflows?

No procedures or workflows - 0 points
Understood/followed but not documented - 0.5
In development or draft form - 1
Documented (without regular revisions) - 1
Documented and reviewed regularly – 1

Q23. Has the institution adopted any particular digital preservation standards, best practices or guidelines? ***or*** *Q26. Has the organization used any digital preservation-related self-audit or gap analysis tools?*

"Yes" in either column counts as 1 point; "No" or "I don't know" as 0.

D. & E. Technology and Infrastructure: Access & Tools

Q29. How does your organization give access to digital materials?

If 'Web platform/repository' selected, 1 point.

Q33. Is your organization creating forensic disk images?

"Yes" counts as 1 point; "No" or "I don't know" as 0.

Q36. Does your organization use digital preservation-related processing tools (e.g. Archivematica, Arkivum, Preservica)?

"Yes" counts as 1 point; "No" or "I don't know" as 0; use counts if in testing or production.

F. Technology and Infrastructure: Storage

Q40. In what storage systems/media are digital assets currently kept? (Select all that apply.)

If at least two managed storage methods selected ("cloud storage," "replicated storage network service," or "tape") - 1 point.

H. Holdings and activities section:

Q58. Approximately how many terabytes of digitized content has your organization created? (Indicate "unknown" if you are unable to ascertain this figure.) ***or*** *Q59. Approximately how many terabytes of born-digital content has your organization collected? (Indicate "unknown" if you are unable to ascertain this figure.)* ***or*** *Q60. If "unknown" was indicated for one or both of the questions above, please indicate the total approximate amount of digital content your organization holds in terabytes. (Indicate "unknown" if you are unable to ascertain this figure.)*

If were able to indicate a figure for at least one column - 1 point.

Q61. Approximately what percentage of born-digital content is stored on each of the following types of storage: (Indicate "unknown" if you are unable to ascertain this figure.)

If were able to indicate a percentage figure adding to 100 - 1 point.

G. Staff & Resources

Q44. List all roles that are responsible/accountable for day-to-day digital preservation activities in your organization and indicate what percentage FTE of each is devoted to digital preservation. (Please estimate to the best of your ability.)

Have at least 100% FTE in total staff listed - 1 point.

Q45. Is your organization intending to expand staff responsibilities for digital preservation?

“Yes” counts as 1 point; “No” or “I don’t know” as 0.

Q49. Approximately what percentage of your organization’s last completed fiscal year expenditures were dedicated to digital preservation (including salaries, storage costs, tools and technologies, etc.)? (Indicate "unknown" if you are unable to ascertain this figure.) ***or*** *Q50. Do you anticipate increasing expenditures dedicated to digital preservation in the next 1-2 years? - By what anticipated percentage? (Indicate "unknown" if you are unable to ascertain this figure.)*

Either indicated percentage value in question 6 or said “Yes” to Q50 **and** were able to indicate the anticipated percentage increase in question 7 - 1 point.

Appendix 3: Q18 - Digital Preservation-related organizations, Conferences, Projects and Initiatives - Word Cloud Acronym Key and Weights

The following is a list of organizations, conferences, and projects and initiatives that survey respondents listed when asked "Is your organization a member of, or participant in, any regional, national, or international organizations, conferences or projects specific to digital preservation?" Distinct entities were counted once per mention per respondent. If a respondent mentioned several projects or initiatives being organized by the same umbrella organization, such as the CARL Portage Network and CARL's Canadian Web Archiving Coalition, then this was counted once in Table 1. However, see Table 2 for a listing and total of all mentioned sub-projects or initiatives.

Table 1. List of main entities by number of mentions, and then alphabetically.

Mentions	Acronym	Full Name
13	CARL	Canadian Association of Research Libraries
9	COPPUL	Council of Prairie and Pacific University Libraries
9	OCUL	Ontario Council of University Libraries and Scholars Portal
6	iPRES	iPRES International Conference on Digital Preservation
5	CRKN	Canadian Research Knowledge Network
5	Portico	Portico
4	ACA	Association of Canadian Archivists
4	DDPDG	Canadian Heritage Information Network Digitization and Digital Preservation Discussion Group
4	NHDS	National Heritage Digitization Strategy
4	PKP-PLN	Public Knowledge Project Private LOCKSS Network
3	CAUL	Council of Atlantic University Libraries
3	CGI-PLN	Canadian Government Information Private LOCKSS Network
3	CLOCKSS	Controlled Lots of Copies Keep Stuff Safe
3	LOCKSS	Lots of Copies Keep Stuff Safe
2	AABC	Archives Association of British Columbia
2	BitCurator	BitCurator Consortium

2	CCA-ACPDS	Canadian Council of Archives Archives Canada Digital Preservation Service
2	COAR	Coalition of Open Access Repositories
2	CPTA	Council of the Provincial and Territorial Archivists
2	DLF	Digital Library Federation
2	HathiTrust	HathiTrust
2	IIPC	International Internet Preservation Consortium
2	Internet Archive	Internet Archive
2	NPTAC	National, Provincial and Territorial Archivists Conference
2	PASIG	Preservation and Archiving Special Interest Group
2	SAA	Society of American Archivists
2	USGD-PLN	United States Government Documents Private LOCKSS Network
1	AAO	Archives Association of Ontario
1	ARLIS	Art Libraries Society of North America
1	ARMA	Association of Records Managers and Administrators
1	ASA	Archives Society of Alberta
1	BCDL	British Columbia Digital Library
1	BCI	Bureau de coopération interuniversitaire
1	CCI	Canadian Conservation Institute
1	CDFC	Canadian Digital Forensics Community
1	CNI	Coalition for Networked Information
1	DPC	Digital Preservation Coalition
1	DPN	Digital Preservation Network
1	DuraSpace	DuraSpace
1	FPTRC	Federal-Provincial-Territorial Records Council
1	ICAM	International Confederation of Architectural Museums
1	InterPARES	InterPARES

1	IRLA	Independent Research Libraries Association
1	Islandora	Islandora Foundation
1	Lyrasis	Lyrasis
1	OMA	Ontario Museums Association
1	OPF	Open Preservation Foundation
1	SAFE-PLN	SAFE Private LOCKSS Network
1	SPN	Software Preservation Network

Table 2. Sub-entities of above organizations, organized by name and total number of mentions

Mentions	Acronym	Full Name
	ACA	**Association of Canadian Archivists**
1	TAATU	The Archives and Technology Unconference
	BCI	**Bureau de coopération interuniversitaire**
1	Geoindex	GeoIndex
	CARL	**Canadian Association of Research Libraries**
8	CWAC	CARL Canadian Web Archiving Coalition
6	Portage	Portage Network (general)
4	DPWG	CARL Digital Preservation Network
2	DVN	Portage Dataverse North Working Group
1	PEG	Portage Network Preservation Expert Group
	CAUL	**Council of Atlantic University Libraries**
3	DSPC	CAUL Digital Preservation and Stewardship Committee

	COPPUL	**Council of Prairie and Pacific University Libraries**
5	DSN	Digital Stewardship Network
4	LOCKSS/WestVault	COPPUL Private LOCKSS Network/WestVault
2	DSNSC	COPPUL Digital Stewardship Network Steering Committee
	CRKN	**Canadian Research Knowledge Network**
2	Canadiana	Canadiana
2	PAC	CRKN Preservation and Access Committee
2	TDRTG	CRKN Trusted Digital Repository Task Group
	DuraSpace	**DuraSpace**
1	Fedora	Fedora
	OCUL	**Ontario Council of University Libraries and Scholars Portal**
5	Permafrost	Permafrost
2	OLRC	Ontario Library Research Cloud
3	DCC	Digital Curation Community
2	Dataverse	Dataverse
1	Canopus	Canopus
		Various host organizations
3	DPM	Digital Preservation Management Workshop

Appendix 4: Q31 - Other Access Tools/Repositories in Use

Table 1: List of access tools/repositories under free-text "other" option in Q31.

System	Digital objects only*	Descriptions only	Digital Objects and Descriptions	Not specified
Argus			1	
Artstor Shared Shelf			1	
Artudis				1
Avalon Media System			1	
Axiell Emu			1	
CollectionSpace		1		
Drupal front end			2	1
Eloquent WebGENCAT			1	
EPrints			1	
Ex Libris DigiTool			1	
InMagic		1	2	
Omeka			2	1
Preservica	1			
The Museum System		2		

Appendix 5: Q33 - Digital Forensics-related Tools in Use

Table 1: Full list of forensics-related tools under free-text option in Q33.

Tool	Number of mentions
BitCurator	10
FTK Imager	4
Kryoflux	4
Tableau write-blocker	3
Floppy drive	2
FRED	2
Guymager	2
ISO Buster	2
AFF (Advanced Forensics Format)	1
Anadisk	1
Atari800 emulator	1
AtariWriter	1
EWF (Expert Witness Format)	1
FC 5025 controller	1
ImgBurn	1
Iromlab	1
R-Studio	1
Tape drives	1
VirtualBox	1

Appendix 6: Notes on Areas for Survey Improvement

The following are areas where the survey could be improved:

- The survey did not capture potentially more granular states of policy/procedure development in questions 20 and 22. Many organizations may have individual policies or procedures at different levels of development, while the survey required them to generalize the status of these documents.
- As an alternative to organizing questions 36-38 (preservation processing tools) around the use of specific tools, the survey could also organize questions around preservation functions, such as fixity checks, file format identification, etc. for a different picture of how often these functions are being performed. However, a related issue to the above is capturing how granular these functions might be over an organization's entire set of collections.
- In response to question 57, many organizations do not prioritize digitization projects based on format. This question might be reformatted or omitted from future surveys.
- In relation to question 63, respondents could have been asked to assess themselves on the NDSA Levels of Preservation instead of the maturity scale.