Strategy for a Digital Preservation Program

Library and Archives Canada

November 2017

Table of Contents

1. Introduction

Library and Archives Canada (LAC) acquires, manages, preserves and provides enduring access to Canada's documentary heritage, and serves as the permanent repository of records of the Government of Canada, as well as publications and records of historical or archival value. LAC is the sole organization in the Government of Canada with a national mandate for long-term preservation. This mandate is enshrined in LAC's legislation, empowering the Librarian and Archivist to ensure that documentary heritage is available to Canadians for the long term.

LAC's library and archival collections contain a wide range of textual, visual, audiovisual, and web content that is supported by a variety of software, hardware and operating systems in a multitude of formats that are vulnerable to technological obsolescence, media deterioration and decay. The LAC Digital Archive at the Preservation Centre serves as the repository of LAC's permanent digital collections. The Preservation Centre contains only a fraction of the digital collections that LAC has acquired over the years. Much more needs to be done to ensure that all digital documentary heritage collections are preserved in a manner that meets the LAC mandate.

More than ten petabytes of digital material are contained on Linear Tape Open (LTO) tape in the Preservation Centre storage vaults. The vaults contain primarily digitized and audiovisual holdings.

This document outlines how LAC intends to develop a fully operational digital preservation program.

2. Definition and scope

Digital preservation is defined as "the active management of digital content over time to ensure ongoing access."[1] To ensure ongoing access, digital preservation staff must proactively monitor and intervene to protect digital heritage content from technological obsolescence.

LAC is aligned with the international standard ISO 14721:2012, Reference Model for an Open Archival Information System. The OAIS reference model describes the functions and roles of a digital archive and helps define the key elements needed to put in place a digital preservation program. LAC uses the OAIS reference model to define the scope of its digital preservation program.

Based on the definition of the Digital Curation Centre (DCC) Lifecycle Model, digital preservation at LAC is one part of the wider set of functions in the digital lifecycle. The functions we are addressing in the DCC model are: ingest, preservation action, store and transform. Addressing the full digital lifecycle is outside the scope of this strategy. Digital acquisitions and digital discovery are also out of the scope of this strategy and will be considered separately in future LAC strategy documents. Although digital preservation is not directly responsible for the

---

[1] Source: http://www.digitalpreservation.gov/about/

full digital lifecycle, we will continue to provide guidance and advice to colleagues to ensure the integration of the various components.
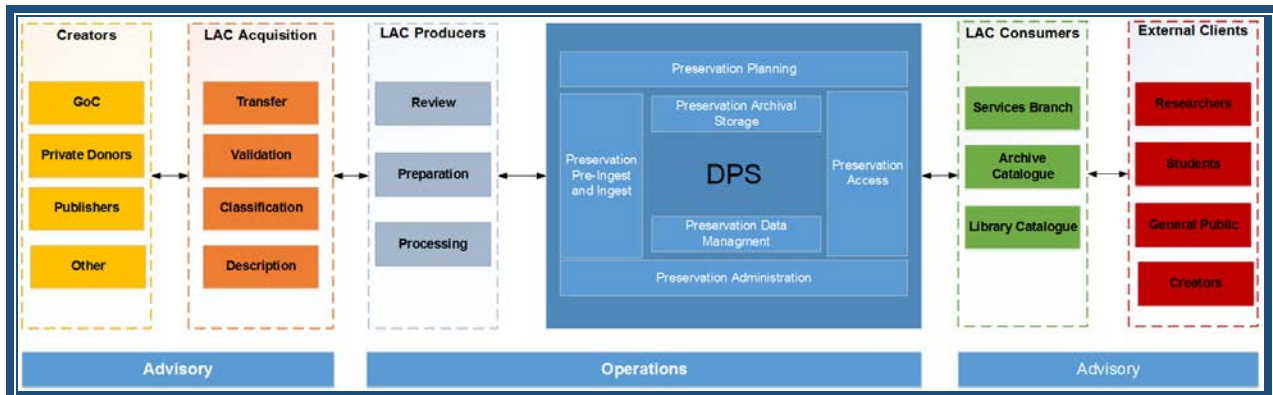

Figure 1: Illustration of organizational scope and positioning of the Digital Preservation Solution (DPS) at LAC

This new strategy for the digital preservation program partially supports LAC's 2015 digital strategy, and specifically the digital curation pillar. The 2015 digital strategy states that "LAC will truly be a trusted digital repository…." One of the outcomes of the digital curation platform in the 2015 strategy is that LAC is aligned with industry standards for trusted digital repositories.

3. Vision for digital preservation

By 2024, LAC will have a sustainable digital preservation program that is compliant with ISO 16363: Audit and Certification of Trustworthy Digital Repositories. ISO 16363 is an international standard that is based on ISO 14721. It is a checklist of deliverables for a digital preservation program, and provides a set of criteria that organizations must meet to implement a trusted digital repository.

The enabling factors that LAC must improve upon in order to achieve its goal for a digital preservation program are outlined in the following end-state vision:

End-state vision for the Digital Preservation Program

The ideal end-state is for LAC to have an internationally recognized trusted digital repository program that has all the functions of the OAIS implemented, including technology watches, integrity checking and preservation planning.
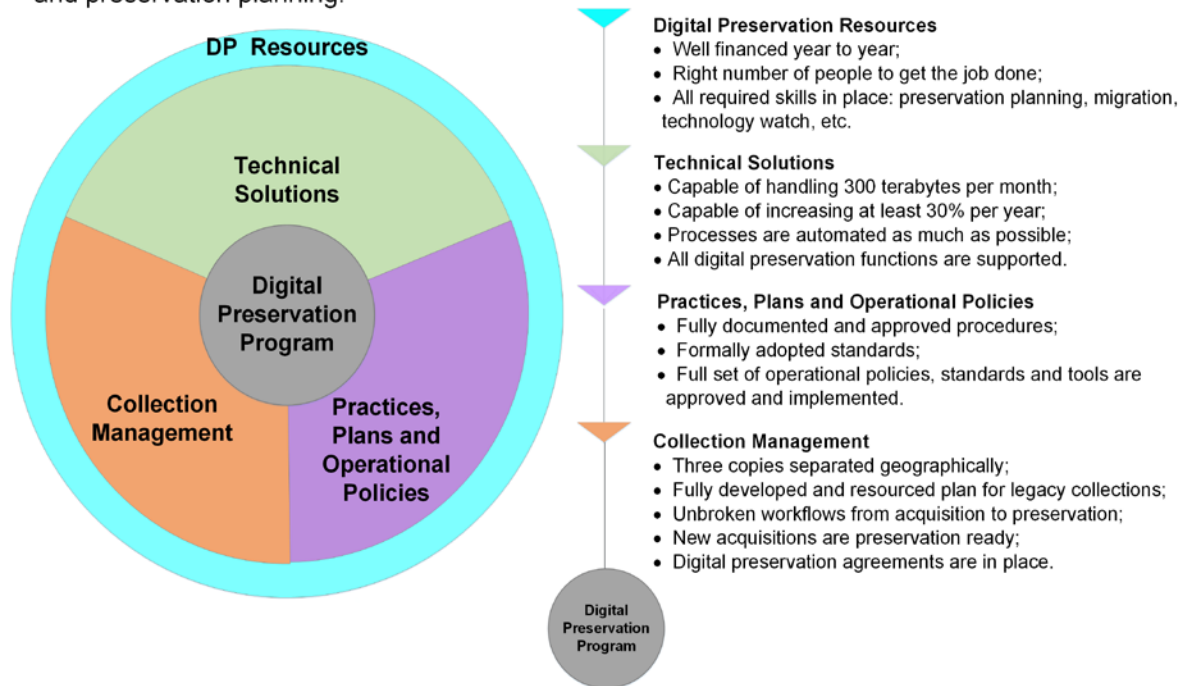
Digital Preservation Resources
- Well financed year to year;
- Right number of people to get the job done;
- All required skills in place: preservation planning, migration, technology watch, etc.

Technical Solutions
- Capable of handling 300 terabytes per month;
- Capable of increasing at least 30% per year;
- Processes are automated as much as possible;
- All digital preservation functions are supported.

Practices, Plans and Operational Policies
- Fully documented and approved procedures;
- Formally adopted standards;
- Full set of operational policies, standards and tools are approved and implemented.

Collection Management
- Three copies separated geographically;
- Fully developed and resourced plan for legacy collections;
- Unbroken workflows from acquisition to preservation;
- New acquisitions are preservation ready;
- Digital preservation agreements are in place.

Figure 2: Illustration of the end state vision

The crucial elements in the development of a digital preservation program are defined in ISO 16363. To deliver on this vision, LAC will:

- ➢ Implement sustainable technical solutions
- ➢ Establish a collections management framework to ensure that digital collections are systematically preserved
- ➢ Produce evidence of program development, which is a key requirement for ISO 16363 certification
- ➢ Design and implement the organizational framework for a program while streamlining business processes
- ➢ Secure adequate human and financial resources

As we begin to shape the future of digital preservation at LAC, this strategy provides the framework to manage and coordinate the activities needed to bring the organization to progressive levels of maturity. Each phase is broken down into a set of interrelated deliverables and senior executive checkpoints to validate the soundness of the approach.
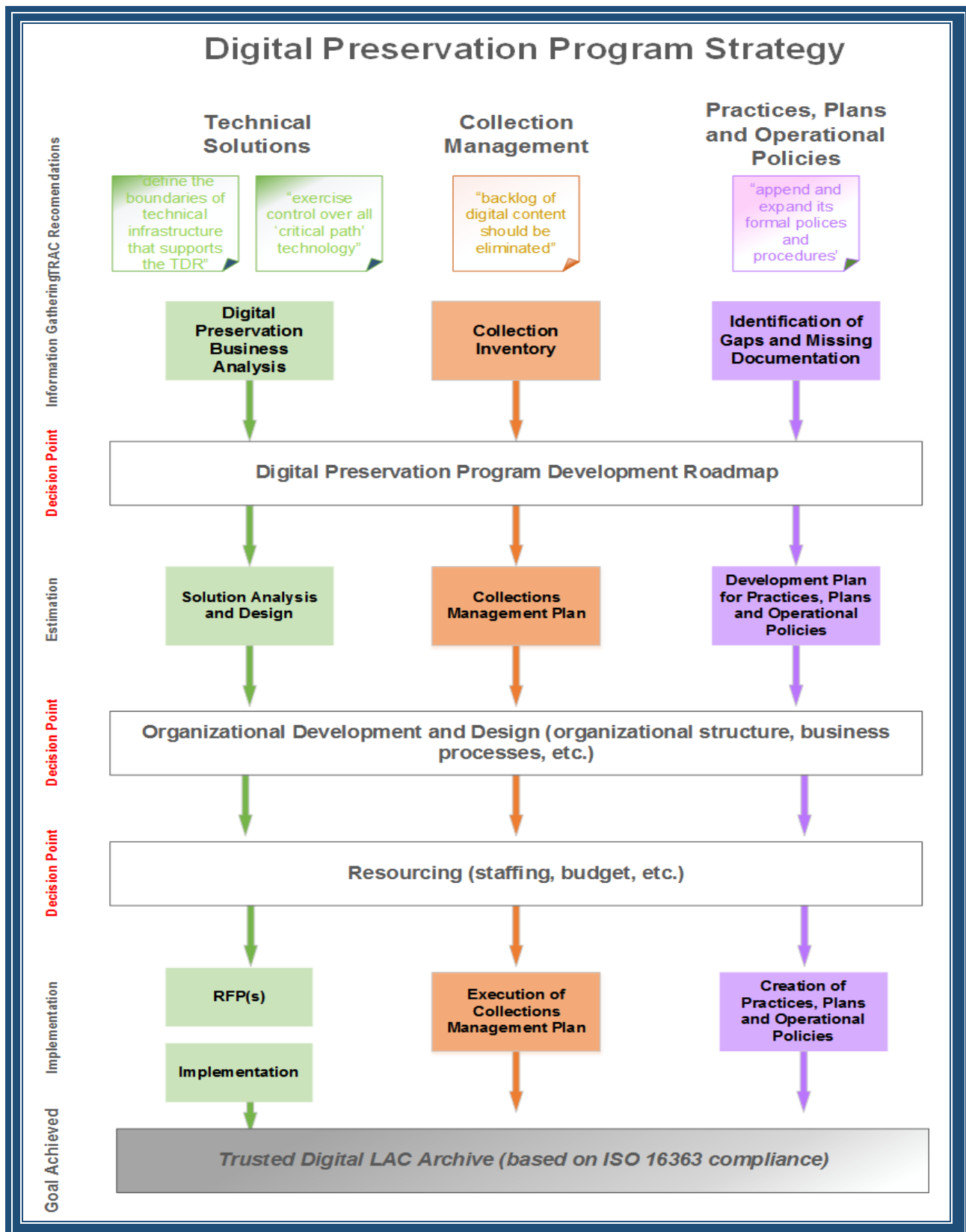
Figure 3: Illustration of the overarching framework for a sustainable digital preservation program.

The section below outlines the steps that LAC must take to overcome its digital preservation challenges.

3.1 Phase 1: Information gathering

This phase embodies the program foundation as the baseline for development. Main activities will be to explore, identify and document digital preservation business needs and issues, a fundamental exercise in problem definition and program specification. The results of this phase will be to identify the size and extent of LAC digital preservation challenges, define the technology needs, and complete a gap analysis of the policy documentation in order to meet ISO 16363 requirements.

Digital preservation business requirements, which were completed in June 2017, will be used as the starting point of discussions to streamline business processes and develop the business technology solutions to support long-term preservation of digital holdings. This document provides information on the size and scope of the technology required and serves as the catalyst for the evolution of the program. It will also provide senior managers with information to make investment decisions on the digital preservation infrastructure.

As part of the Digital Collections Inventory, a suite of collection profiles was created to get a better understanding of digital collections acquired by LAC since the 1970s. The profiles allow LAC to:

> ➢ Quantify the digital holdings including backlogs of collection to be preserved
> ➢ Identify who is responsible for them
> ➢ Discover where they are located
> ➢ Define their file formats and other collection characteristics
> ➢ Determine whether or not they are preserved
> ➢ Assess their risk of permanent loss.

These collection profiles describe the scale of the work required to preserve all digital holdings and will help guide future preservation planning.

A policy gap analysis will be conducted during this phase to identify the operational plans, procedures and practices to be prioritized for development. This analysis will guide LAC in establishing a formal policy framework for the preservation of Canadian digital documentary heritage.

3.1.1 Success factors of Phase 1

One of the key outcomes of Phase 1 is the development of a roadmap, which will provide direction for future program development and implementation. It will describe planned steps to

build digital preservation capacity. This document will lay out major activities to be completed, roles and responsibilities, the sequence of their execution and general timeframes.

Phase 1: Information gathering	Purpose	Action items / deliverables
Technical Solutions	To identify features needed in a system and delineate the size and extent of the problem.	Digital preservation business requirements and an options analysis.
Digital Collection Management	To determine the size and nature of the digital collections and assess for risk of permanent loss.	The digital collections inventory for published heritage, private archives, and government records.
Practices, Plans and Operational Policies	To identify gaps in policy instruments and associated documentation as a first step to meeting the ISO 16363 requirement for documentary evidence.	A gap analysis of digital preservation policy instruments.

Decision Point: end of Phase 1	A **roadmap for a sustainable digital preservation program** will help us arrive at time and cost estimates for the remainder of the strategy.

3.2 Phase 2: Program development

This phase aims to obtain further precision on the estimates obtained in Phase 1 and to frame what the program will look like. By the time we get to Phase 2, the size and extent of the problem will be better understood. The next steps will be to design and deliver digital preservation capability.

Technical solutions: Based on the digital preservation business requirements completed in Phase 1, this phase will determine the appropriate technologies and infrastructure in a digital preservation solution, and will focus on constructing the solutions analysis and design. It will include technical design aspects for storage infrastructure, networks, platforms and applications as well as interoperability and integration with other LAC systems. The cost and effort for a new business solution will be analysed and options presented for a management decision.

Building on the digital collections inventory, LAC will develop a digital collections management plan to assign preservation processing priorities for the backlog of digital collections, requirements for their preservation, and the level of investment required for each collection. One of the key activities of this phase will be to conduct a more detailed sampling of the collections to assess for data corruption and loss. Technical specifications and standards for preserving the various types of content will also need to be developed.

A practices, procedures and policies development plan will be issued in Phase 2, based on the gap analysis completed in Phase 1. Policy decisions that have cost implications, such as number of copies, storage and diversity of geographic locations, will be identified. Workflows, processes and procedures will be further defined to allow material to flow seamlessly from acquisition to preservation and discovery.

3.2.1 Success factors of Phase 2

At the end of Phase 2, LAC will have achieved clarity in its program directions. LAC will have defined what constitutes scalable services and infrastructure for the program. The level of investment for the technology, infrastructure and resources required will have been determined, as well as costing models. A knowledge base of the extent of digital content to be preserved, levels of staffing required, collection priorities, new collections to be preserved, costing options to deliver the program, and appropriate budget levels assigned will inform this phase.

Phase 2: Program development	Purpose	Action items / deliverables
Technical Solutions	To design potential technological solutions based on business requirements developed in Phase 1 To conduct proofs of concept To assess technology investment required	Technical solution analysis and design to define the appropriate hardware and software technologies and other components needed for long-term preservation
Digital Collection Management	To prioritize preservation of existing collections and incorporate planning for new collections	Digital collections management plan

Practices, Plans and Operational Policies	To develop required policies and lower-level instruments	Development plan for practices, plans and operational policies

Decision Point: end of Phase 2	1. **Organizational development and design.** One of the deliverables of this phase will be a detailed organizational plan identifying program and organizational capacities, including skills and competencies required for a digital preservation workforce.
	2. **Financial and human resources.** LAC will secure sustainable resources for the program.

3.3 Phase 3: Program implementation

LAC will be at a mature state of readiness to start program implementation, as it will have a shared vision, institutional commitments, management support, organizational capacities, consistent workflows and dedicated resources assigned to the program.

By the end of Phase 3, LAC will have achieved its goal for a sustainable digital preservation program and a trusted digital repository in accordance with national and international best practices and industry standards.

3.3.1 Success factors of Phase 3

For the technical infrastructure, a technical design will have been completed, and LAC will be equipped with fully formed business information with which to implement the design. The result of this Phase will be a fully functioning technology solution to support digital preservation.

Based on decisions made in the collections management plan in Phase 2, LAC will be equipped with the necessary decision tools and infrastructure to execute collection preservation priorities. The collections management document will need to be refreshed regularly as more collections are acquired. As digital collections management is implemented, the collections management plan becomes an essential part of preservation planning and administrative workflows.

LAC will develop and implement digital preservation policy instruments that are integrated into the business of the organization. It will also implement a regular schedule for policy and program evaluation as part of an ongoing review cycle.

LAC will conduct an ISO 16363, Audit and Certification of Trustworthy Digital Repositories, formally or informally, to evaluate the progress of the program implementation. The organization's approach to audit and certification will be determined during this phase. Given the

vast extent of digital collections at LAC and the level of documentation required, full formal certification with a third party auditor that has been accredited against ISO 16919 could be a costly and time-consuming proposition. Other approaches may involve informal certification through peer review, or through self-assessment. LAC may also choose to audit its trustworthiness one collection at a time.

Phase 3: Program implementation	Purpose	Action items / deliverables
Technical Solutions	To implement business solutions.	Issue RFP(s) for a digital preservation solution.
		Implement a scalable digital preservation solution.
Digital Collection Management	To address collection preservation priorities.	Execute collection management plans such that current and new collections are preserved in the LAC Digital Archive.
Practices, Plans and Operational Policies	To produce documentation to meet ISO 16363 documentary requirements.	Creation of practices, plans and operational policies, with ongoing updates.
		Determine the level of ISO certification as one of the decision points.
		Documentary evidence publicly posted to comply with ISO 16363 requirements.
		Program evaluation to ensure goals of the program are being met.

The digital preservation program strategy will be refreshed as deliverables for each phase are completed, or as the digital preservation program crystallizes and becomes integrated into the fabric of the organization.

4. Conclusion

LAC's strategic approach to digital preservation will need to evolve dynamically as the program matures. Preservation plans and services must be regularly reviewed and updated for continued alignment with the organizational mandate and strategic objectives. Perpetual service improvements to address legacy and new digital content must be part of the organizational planning cycles, rationalized against technical and other service provision models in the digital ecosystem.

It is paramount that LAC demonstrate constancy of purpose and dedication to its digital preservation obligations to ensure the survival of the nation's documentary heritage. The consequences and cost of inaction would be LAC's failure to meet its legislated mandate and obligations to Canada's citizens and the loss of national digital heritage content.