@Risk North 3 Summit Report

Gatineau, Québec | November 21-22, 2024

February 2025

CARL Digital Preservation Working Group

Co-organizers of the @Risk North 3 Summit

www.carl-abrc.ca

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

Table of Contents

Executive Summary

The @Risk North 3 summit (Gatineau, November 2024), organized by the CARL Digital Preservation Working Group (DPWG) and partners, addressed current challenges in the digital preservation arena through presentations by key stakeholders, guided discussions and lightning talks. Areas under discussion included broad themes such as education and training, people strategy, funding and resource allocation, and more focused ones like web archiving, research data, benchmarking, and tools and technologies.

Across the summit, participants emphasized the need for continued and deepened collaboration between institutions, organizations and practitioners for the purposes of advocacy, knowledge-sharing and skills development. Identifying sustainable solutions, shared infrastructure among them, addressing growing storage requirements and budgetary constraints was a priority. The community could benefit from establishing and sharing best practices, methodologies and workflows around key activities such as benchmarking, and greater mechanisms to facilitate access to resources. Varied and more extensive training, both formal and informal and at all career stages, would benefit individual practitioners and organizations.

Moving forward, a follow-up webinar to share findings, the creation of a community of practice, and planning for future iterations of the @Risk North summit would continue the momentum and encourage the relationship-building the summit fostered. Further, the DPWG will work towards developing a multi-year action plan, implementing a national benchmarking exercise, and creating a CARL Visiting Program Officer position in Digital Preservation to aid in coordination and planning.

Introduction

On November 21-22, 2024, 120 leaders and professionals working in digital preservation, libraries, archives, and related organizations convened at the Library and Archives Canada (LAC) campus in Gatineau, Québec, for @Risk North 3 (@RN3). The third iteration of this Canadian digital preservation summit (previous summits took place in 2017 and 2018) was organized by the Canadian Association of Research Libraries (CARL) Digital Preservation Working Group (DPWG), in partnership with LAC, the Canadian Research Knowledge Network (CRKN), Bibliothèque et Archives nationales du Québec (BAnQ), Internet Archive Canada (IAC), the Digital Research Alliance of Canada (Alliance), and the Digital Preservation Coalition (DPC). The DPWG represents the collective effort of CARL and a small number of non-CARL members to assist Canadian research libraries and adjacent organizations in navigating the significant challenges of their evolving digital stewardship roles.

The DPWG organized the in-person summit to enhance digital preservation strategies both locally and nationally, increase collaboration at scale, and to better utilize shared infrastructure among research organizations. A further goal was to emphasize digital preservation within broader digital strategy and innovation plans, as well as the significance of sustained long-term stewardship amidst pressures for immediate return on investment. Collaboration is a strong characteristic of Canadian library work, with numerous examples found at both regional and national levels. @RN3 aimed to expand on that ethos and those existing relationships to actively do digital preservation – an ongoing cycle of risk identification, mitigation, and corresponding actions.

The @RN3 program featured reports on the current state of digital preservation both in Canada and internationally, with representatives from organizations such as CRKN, Scholars Portal, Internet Archive, and DPC, as well as larger institutions like BAnQ and LAC, along with smaller academic and cultural organizations. This vital work was shared through rich sets of panel discussions, roundtables, and lightning talks. Summit organizers aimed to make @RN3 a "working meeting," carefully designing the program to maximize opportunities for table talk and collegial discussion, limited by the time constraints of the one-and-a-half-day meeting.

1. Information Sessions: Presentation Summaries

1.1. Opening Keynote: Global Perspectives and the Evolution of Digital Preservation Work, William Kilbride (DPC)

In his keynote address, William Kilbride, Executive Director of the Digital Preservation Coalition (DPC) enumerated current challenges in digital preservation work, including storage, standards, technical debt and dependency, and underscored the importance of trust and collaboration in mitigating these challenges. While, for example, the maturity of approaches, services, policy and funding are important, taking collaborative approaches to problem solving, establishing competencies and creating new processes and technologies is essential. Collaboration occurs on multiple fronts, between individuals and agencies, and across generations and nations. Kilbride stressed the importance of shared infrastructure as one such collaborative approach.

1.2. National Service Providers

Susan Haigh (CARL) moderated a panel of representatives from national digital preservation service providers, in which Jason Friedman (CRKN), Kate Davis (Scholars Portal), and Andrea Mills (Internet Archive Canada) provided updates on the activities of their respective organizations.

Speakers emphasized the strength of community and the importance of collaboration with respect to a growing need for shared services and infrastructure. Larger institutions and organizations working with smaller ones combats isolation faced by those working at smaller institutions and benefits everyone. Even larger organizations are not as large as we might think, and they too rely on community and the dedication of those who do the work on a daily basis. Key tenets, such as the fact that preservation and access are related but distinct, were underscored.

1.3. Digital Preservation in Canada: Workflows, Experiences, Insights

Anna Perricci (DPC) led a discussion about preservation of digital documentary heritage with Kyle Browness (LAC), Pascale Montmartin (BAnQ), and Émilie Fortin (Université Laval).

Speakers emphasized the importance of collaboration to create community and learn from others. For example, creating strong relationships with IT is especially valuable for practitioners, even where there are challenges in navigating respective needs and perspectives. Speakers also shared insights on the decision making involved in selecting digital preservation services and tools. In terms of storage, institutions weighed benefits and detractions of cloud storage. Cloud storage offers enhanced access and tech diversification on the one hand, and introduces risks associated with vendor lock-in, third-party integrity and ongoing need for local storage as back-up on the other. Cost savings may not be significant, and there are technical and environmental sustainability factors to consider. Migration is a time-consuming and

often complicated endeavour. Speakers underscored the need to distinguish between access and preservation, and the fact that there are different guidelines to respect within and across organizations.

1.4. Lightning Talks

Sixteen lightning talks were delivered as part of two sessions in which practitioners shared three-to-five-minute presentations on current projects, innovative approaches, and lessons learned in the Canadian digital preservation context. Brief abstracts of the lightning talks have been compiled in Appendix 1.

2. Discussion Sessions

On both days of the summit, participants joined concurrent breakout sessions for guided discussions on a range of topics.

Discussion Sessions, Part I – Gaps and Needs Identified

2.1. Education and Training

In this session, the first of two focused on education and training, session convenors William Kilbride (DPC), Vivianne Maréchal (UQAM), and Robyn Stobbs (Athabasca) initiated breakout discussions guided by a set of questions around priorities, gaps and needs, partnerships, and resources.

Priorities

Participants viewed network-building, strengthening capacity and competencies, and supporting the growth of practitioners from novice to intermediate as priorities. In terms of identifying which skills should be prioritized, participants emphasized that soft skills, such as problem solving and communication, are as valuable as technical skills. Developing confidence in practitioners and encouraging their ability to make decisions or ask for help is important; this confidence can be nurtured through hands-on experience. Educating administrators on the importance of digital preservation programs is also strategically important. In institutions without a dedicated digital preservation program, having an experienced leader in place to both galvanize interest and build skills in staff is valuable. Similarly, a "digital preservation champion" could advocate within the institution and educate administrators and IT staff. Having a clear strategic vision for digital preservation in Canada is important, as is engagement and leadership from archival institutions, provincial or national.

Gaps and Needs

Digital preservation training should be more accessible, widely available, and offered for different practice areas and levels of expertise. Increased learning in many forms was deemed beneficial, including basic hands-on, staff-wide training in digital preservation, and more extensive modular training in all areas for existing staff. Specific courses at the university level would enhance competencies. Training is limited to on-the-job, or DPC offerings. Tools and training materials should be translated into French. Practitioners need funded professional development opportunities, sustained training, and training that includes follow-up. This goes hand in hand with increased professional development opportunities where a network or cohort can be formed through training opportunities. The Digital Preservation Management Workshops series (MIT Libraries) is viewed as a strong offering. Training in the ethics of the profession, sustainability and environmental impacts, and what to keep or not keep, were raised as areas to further develop.

Partnerships

In terms of partnerships, developing a digital preservation methodology and contextualizing local practices within larger developments and trends in the field were seen as useful strategies, as well as developing bilingual resources and tools. Effective partnerships could also be leveraged through a community of practice, practice exchanges/interchange, student internships, training opportunities from a national institution, IT partnerships, shadowing and mentorship within and between institutions. Inter-institutional mentorship would be particularly helpful for small institutions, where practitioners often experience isolation. OCUL or CARL should help smaller institutions without DPC membership. Canada has a strong archival presence, and this should be more effectively leveraged.

Resources

More training resources would be helpful. Participants shared training offerings and professional development opportunities such as Digitization Working Group (CHIN, War Museum), DPC (advocacy, toolkit for hiring, self-directed and adaptable offerings), University of Ottawa course (but very little otherwise in terms of university-level digital preservation courses), and the Digital Archives Specialist curriculum.

2.2. Web Archiving

This session, led by Geoff Harder (UAlberta; Chair, DPWG), with Tom Smyth (LAC), Mireille Laforce (BAnQ) and Jefferson Bailey (Internet Archive), focused on collaboration opportunities within the Canadian web archiving community, with a broader vision for collaboration across the Western hemisphere.

The speakers each provided an update on web archiving efforts at their respective institutions, describing the scope of their collections and identifying key issues in the field. This was followed by a discussion of the Canadian web archiving community, inter-institutional coordination, shared goals, and the identification of neglected aspects of web archiving that require greater attention.

Institutional Updates

Tom Smyth reported that web archiving is a formally established program at LAC, with a collection of approximately 840 TB, with 200 TB from the 2024 .ca domain crawl. New collections are added on an ongoing basis. Legislative authority held by LAC may give it broader powers to collect resources from the web compared to other institutions.

Mireille Laforce emphasized that BAnQ focuses on preserving Québec's digital heritage through its web archiving program. Varying levels of access create challenges: government and licensed sites are available online to the public, other restricted sites are available onsite only, and others still are collected for preservation purposes with no public access. Login requirements, metered access and blocked crawlers, as well as changing URLs and restructuring of government websites create more specific challenges; multiple tools incur technical challenges.

Jefferson Bailey provided an overview of Archive-It and its scope, noting it serves over 1200 institutional users in more than 40 countries and has archived 4.5 PB of web content with an annual crawl capacity of 700 TB and 3.5 million seeds in the system. Archive-It is used by 106 Canadian organizations, 50 of which participate in Community Webs programs. Internet Archive has also collaborated on projects such as LAC's .ca 2024 domain web crawl and access portal, with tailored features. Challenges include: limited staff and resources; constrained budgets and commercial creep into library and archives space; expensive, niche and complicated tools; limited opportunities for knowledge sharing and community building.

Proposed Collaborative Network

In response to the need for more coordinated efforts and opportunities, Jefferson Bailey shared a proposed initiative for a collaborative network for web archiving with an emphasis on knowledge sharing, professional development and social activities. While Archive–It could provide some seed funding, resources and staff, with further support from other institutions, this would be an independent community. This project has garnered interest in Canada, the United States, Mexico, South America and the Caribbean. Attendees provided feedback, with some expressing reservations about

joining a new organization and the time commitment involved. A venture like this one would not preclude a local chapter or Canadian community from taking root, nor replace the work of a collaborative collecting program like Collective ART Archive (CARTA). The international scope could benefit Canadian practitioners as, for example, the sustainability of a domestic initiative like Canadian Web Archiving Coalition (CWAC) has been challenging.

Coordination

Further discussion surfaced challenges and areas where greater focus is warranted. Increased inter-organizational coordination was called for to grow awareness of web archiving activities. The Canadian Government Information Digital Preservation Network (CGI-DPN), while helpful, is not exhaustive. CWAC was successful in dividing responsibilities and sharing seed lists (i.e. during elections), suggesting the proposed collaborative network could take this up. A tool, platform or registry could be helpful in this respect.

Access

While archivists may successfully collect web resources, they are not necessarily empowered to provide access. Increased knowledge about legal dimensions and institutional risk was identified as an area for further collaboration.

Discovery

Discovery for web archives needs further attention, as these archives may not be well integrated with existing library systems, leading researchers to rely on Google.

Local Preservation

Local preservation is challenging, as organizations do not have capacity to download to local storage; fewer than a quarter of Archive-It users currently do so. This suggests an urgent need for shared solutions, as many organizations lack infrastructure for local preservation.

Crawling Challenges

Capturing media platforms and dynamic content (i.e. YouTube videos, interactive maps) pose technical challenges that could be mitigated by sharing expertise and troubleshooting.

2.3. Research Data

Amanda Tomé (Alliance) and Julie Shi (Scholars Portal) led this discussion, in which participants shared insight on the preservation of research data, surfacing pain points and good practices with each other. The speakers briefly introduced preservation infrastructure and workflows available through their respective services. This was followed by discussion of key challenges and opportunities in the field.

Researcher Engagement

Participants were asked how well they thought research is being captured in Canada. The relationship with the researcher is key and participants noted the importance of connecting with researchers early in the research process to gather the correct information. Different ways of reaching researchers should be identified and effective means of outreach may vary according to discipline. Service providers can help identify gaps in disciplinary coverage. Some researchers may not realize that they have data to preserve, and preservation models should be in place, irrespective of the discipline. Moreover, researchers do not necessarily understand the research life cycle.

Data Re-use

The extent to which deposited data is re-used is not known. The Federated Research Data Repository (FRDR) does not track how data is used. Number of downloads and IP addresses are taken into account. There is a need to track the impact of data using DOIs; it is hard to promote sharing, if usage cannot be demonstrated. Also, there is an issue with succession and encrypted material; for example, what transpires when the person with the decoding key moves on. As it stands, the researcher is the gatekeeper, but not necessarily in perpetuity. Whether the institution has a role to play needs to be determined.

Dataset Size and Sensitivity

Two problems encountered with repositories are the size of the dataset and the optimal curation and access management of sensitive data, including traditional knowledge and Indigenous data. Large datasets have bit-level preservation only, as preservation systems such as Archivematica are not designed to handle. CERN has a solution for big data preservation and the Alliance would like to hold community discussions to gather suggestions and solutions. Many groups, including Alliance RDM, are working to clarify sensitive data types and terminology, needs, and management processes across the research lifecycle.

Mitigating Challenges

In terms of training and competency, avoiding the use of acronyms and technical jargon helps, as does incorporating descriptions of required steps. It is important to understand how things work, but at an accessible level. The Open Archival Information System (OAIS) is the de facto model but has also proven to be a barrier: it is complicated and can be seen as an unattainable goal. Conversely, there is also a problem with capacity if awareness grows and more researchers deposit data. Greater understanding of the differences between data curation and preservation might help. Increased definition of the extent of responsibility of the service provider and the institution is needed to understand respective roles. In terms of collaboration and support, practitioners should reuse what others have done where possible. Software preservation needs more attention.

Research data preservation efforts could be enhanced with communication tools. For example, the DPC has a business toolkit available to demonstrate the monetary value of doing preservation and which is useful for talking to managers. More outreach to researchers is advised so that researchers better understand what they can and should do independently, without library resources. Preservation efforts should begin earlier in the research process and with long-term thinking; preservation is linked to ensuring long term access. This work would benefit from taking a bird's eye view to apprehend and integrate distributed disciplinary research and understand the role of all stakeholders from larger organizations like the Alliance as well as smaller ones.

Discussion Sessions Part II – Shared Challenges and Solutions

2.4. Funding and Resource Allocation

Susan Haigh and Geoff Harder led this discussion structured around four themes: prioritization and positioning; funding; coordination at national scale; and digital preservation value proposition.

Priorities

In the discussion, participants shared that priorities are determined locally, vis-à-vis records and materials directly tied to institutional mandate; these, most often, are not adequately resourced. Smaller institutions in particular are making decisions based on what free tools and services are on offer.

Funding

Leaders are asked to justify the allocation of resources, but digital preservation may be valued in an abstract way rather than as an immediate and ongoing activity. Advocacy is key, as is the language used to frame digital preservation; for example, it can be effective to include mention of data management and data security. It is not compelling enough to be engaged in this work "for the public good." Preservation, it can be argued, can supersede access in importance because the access question is moot if records have not been preserved, and collections stewarded.

Coordination and Partnerships

Project funding is easier to obtain than ongoing funding for infrastructure. In the university context, finding funds to support and maintain infrastructure is particularly difficult. This challenge mitigates in favour of pooling funds for shared infrastructure to avoid reinventing the wheel at the local level. Still, this infrastructure must receive ongoing support, not one-time investment. For example, libraries should not have to fund preservation infrastructure on their own; the government should also support national infrastructure that is relied upon by researchers for the research it funds. Partnership with the Alliance makes sense with respect to research data; greater consideration of how a partnership could work beyond this area is warranted. Looking to Borealis funding as a model, for example, the Alliance could fund new development and libraries could then fund ongoing operations.

National Mandates

Participants weighed the risks and benefits of national mandates. With national mandates in place, libraries fear losing control of their budget funds as their funds may go to another part of the institution or organization. Libraries do work with the Tri-Agencies, but there is not a lot of funding for mutual goals. Funding streams may be stronger outside of Canada: in Europe, governments fund more actively and in the United States, there are a range of both private and government funders including the Institute of Museum and Library Services (IMLS).

Key Takeaways

Key takeaways of the session include: increase advocacy; move towards collective approaches and strategy while not jeopardizing funding; re-strategize unfunded mandates; articulate value proposition; emphasize preservation as part of the research lifecycle and part of the cost of doing business, and; consider proportional investment.

2.5. Benchmarking

Amanda Tomé led this discussion to identify the benefits and challenges of measuring digital preservation capabilities. Measuring capabilities is an effective way of

managing work, identifying gaps and supporting advocacy. The effective discoverability of a collection is a great way of measuring success and may exceed volume of preservation in terms of its value. Narrowing the scope, volume and extent of preservation may help the sustainability of a digital preservation program.

Discussion revealed a range of institutional approaches to benchmarking, and that those approaches are subjective. Needs may differ between administrators and practitioners. A sustained practice of benchmarking depends on institutional commitment to devote time to carry it out, annually or bi-annually. Policies, available from DPC and CARL, can be valuable, though it is helpful to have some experience first. Benchmarking can pay off in helping to establish digital preservation program priorities, lending added value through National Digital Stewardship Alliance (NDSA) levels, helping justify decisions, and supporting resource requests. Further, benchmarking can help manage institutional expectations and, in this respect, maintaining transparency about the current level of preservation reached should be encouraged.

The discussion concluded more benchmarking should be undertaken at institutions and this be fostered by some webinars and other training mechanisms. The range of institutions and their respective benchmarking needs call for flexible tools.

2.6. Effective Training and Education – People Strategy

In this session, which was the second focused on education and training, Jason Friedman (CRKN) asked participants to consider the current state of training opportunities available and how they can be improved. Often training is offered as a "one-off", sporadically or siloed. Archivists are not as well-represented as librarians. Membership requirements and cost to non-members are a barrier to participation. Barriers to learning also exist around language: francophones may be challenged with tools and training available only in English, and jargon can be confusing to the uninitiated.

Local and Institutional Training

Practitioners would benefit from collaborations between IT and digital preservation specialists, combined training opportunities, peer education (i.e. Library Carpentry style), support to work through exercises, and user-driven rather than vendor-driven opportunities. Peer-led training and having the opportunity to explore and fail is seen as an effective approach.

National Training

Developing a national program collaboratively with all stakeholders was seen as valuable; participants supported the idea of training for everyone, and the importance of connecting the idea of preservation with specific technical learning. Suggested areas for training include: the basics (e.g. a 1-hour overview of archiving, preservation, access and distribution), copyright, and programs that increase adherence to FAIR and OCAP (Ownership, Control, Access, and Possession) Principles. At the intermediate level, developing a Librarian Research Institute-style workshop for digital preservation was suggested, with a working group put in place to develop an intensive course. Taking a step back, archival training offered as part of Library and Information Studies training may be minimal and not up to date; improving student learning would demonstrate to those entering the profession that archival work is inviting, interesting, and important for all information professionals.

An organization, such as CARL, LAC, or the DPC could be tapped to lead the creation of a training program and/or act as a repository for open learning materials and educational opportunities. Comparable examples include the Canadian Conservation Institute (CCI), the Canadian Heritage Information Network (CHIN), and the Centre de conservation du Québec (CCQ). Partners, including the National Indigenous Knowledge and Language Alliance (NIKLA) and the First Nations Information Governance Centre (FNIGC) could be tapped as collaborators. If in-person training opportunities are created they must be accessible to all, and travel must be supported and accessible to all. Consideration should also be given to how any new initiatives could best serve and include Indigenous communities.

2.7. Tools and Technology

In this session, led by Kyle Browness (LAC), participants worked through an extensive questionnaire in small groups. Together, they contributed information and insight on their use of preservation tools and technologies and experiences with training, collaboration, experimentation and AI. Participants shared best practices in terms of maintaining data integrity, working with complex assets, and addressing file format obsolescence, migration, scalability, and metadata capture.

The wide-ranging responses this session generated have been compiled in Appendix 2.

2.8. Table Discussions – Key Challenges

Lisa Miniaci (BAnQ) headed up this session of table discussions where participants synthesized gaps and challenges currently experienced and identified key ideas to carry forward over the next five years. The discussions were wide-ranging, with extensive feedback provided in the areas of technologies, processes and preservation,

and people. Despite the challenges articulated, participants identified many opportunities. These responses have been compiled in Appendix 3.

2.9. Action Planning

A final session held in plenary, chaired by Anna Perricci (DPC), sought to distill key themes gathered over the course of the summit.

Community

A community of practice would help combat the "lonely manager syndrome" and extend the positive networking that has taken place at the summit. A community of practice could provide learning opportunities that would enhance everyone's understanding. Smaller institutions need a seat at the table and want access to the same benefits offered to larger institutions such as, for example, access to DPC resources. Inter-institutional mentorship should be explored. A working group could be created that includes representatives from both small and large institutions; however, participating in a working group may not be an option for an individual from an institution where they are the sole practitioner. A smaller institution should be represented on the DPWG.

Resources and Training

More effective training will raise the bar in the profession. Practitioners would benefit from a common understanding of what records go to what kind of storage; a set of best practices should be developed. Drop-in clinics, offered by the DPC or others, help address commonly encountered problems. The DPC's "Novice to Know-how" (N2KH) provides 15-20 hours of modular learning. Bilingual resources should be enhanced; BAnQ could be tapped to help with translation.

Support for Shared Infrastructure

Greater support for shared storage needs to be nurtured. Developing a shared infrastructure should be prioritized. Scholars Portal's capacity to support national infrastructure services such as Borealis, Scholaris and OLRC (Ontario Library Research Cloud, their cloud storage service), and to include preservation workflows in these, will be strengthened if more institutions join. Given that everyone will have to do more with less, shared models make sense.

Advocacy

Digital preservation must be part of an on-going narrative. There is a need for champions in this area who are able and supported to vocalize challenges and

advocate for solutions. There should be someone in an oversight role, i.e. a paid position, to coordinate activities and spearhead advocacy. Environmental sustainability is an urgent issue; with costing tools available, more information about costs can be communicated to leadership. For some, maintaining focus on United Nations Sustainable Development Goals (SDGs) will necessitate hard choices about personal values and professional obligations.

Finally, in addition to a summit report, participants expressed interest in a follow-up webinar where findings can be shared, a repository for the lightning talks, having summit materials made available on the CARL website, and further planning done to ensure summits of the nature of the @Risk North series continue to take place at more frequent intervals, at least one every three years.

3. Next Steps

The CARL Digital Preservation Working Group (DPWG) is currently identifying priorities and work plans to build on the recommendations, directions, and momentum that emerged from @Risk North 3.

Several potential actions have been noted, such as the production and public distribution of this final report. As follow-up, the DPWG will also be examining:

- A multi-year action plan to align CARL's work with that of the larger community, fostering collaborative engagement in national planning and meaningful achievements in digital preservation and stewardship.
- A national benchmarking exercise using DPC's Rapid Assessment Model to gain insight into the status of digital preservation activities among Canadian organizations, helping us better understand the opportunities and challenges for capacity building and support for improvement.
- The potential creation of a CARL Visiting Program Officer in Digital Preservation to assist with coordination, community development, training, and communications.

CARL and partner organizations, LAC, CRKN, BAnQ, the Alliance, DPC and Internet Archive Canada will continue to advocate for digital preservation, recognizing that numerous organizations across varied sectors, with diverse professional expertise and perspectives, are essential for making progress toward our goals.

Finally, thank you to everyone in the community who participated in @Risk North 3, whether through survey input, planning, event attendance, or other means. We extend

our gratitude to all those who continue to collaborate to address the challenges we face in preserving scholarly and cultural digital material.

The CARL DPWG is a joint working group, launched in 2016, currently made up of representatives from the following organizations: Canadian Association of Research Libraries (CARL), Bibliothèque et Archives nationales du Québec (BAnQ), Canadiana.org/Canadian Research Knowledge Network (CRKN), Council of Prairie and Pacific University Libraries (COPPUL), Library and Archives Canada (LAC), University of Alberta, University of Toronto, and University of Victoria.

Acknowledgements

Many hands were involved in the planning for @Risk North 3. From start to finish, it was a full team effort to make the event possible, with countless hours of work volunteered and contributed by committed colleagues from many different organizations.

As Chair, I want to thank our sponsors and host organizations: Canadian Association of Research Libraries, Canadian Research Knowledge Network, Digital Research Alliance of Canada, Digital Preservation Coalition, Internet Archive Canada, BAnQ, and Library and Archives Canada. The event would not have been possible without your support and contributions to planning, including the amazing onsite facilities and logistics planning supported by LAC's Veronica Berry, and facilitation activities undertaken by DPC's Anna Perricci.

CARL's Digital Preservation Working Group realized an opportunity and ran with it, so my thanks to the members of the working group, with special acknowledgement to the planning sub-committee: Kyle Browness (LAC), Émilie Fortin (Laval), Jason Friedman (CRKN), Steve Marks (Toronto), Lisa Miniaci (BAnQ), and Amanda Tomé (Alliance). Your expertise, insight, and commitment to the success of the summit was truly incredible.

Deep appreciation and thanks are warranted to the staff at CARL, including Elizabeth Kalbfleisch, who did much of the synthesis and writing for this report, Katherine McColgan, Houda Tarib, Taleen Aktorosian, and unwavering support and guidance from Susan Haigh.

Finally, on behalf of the DPWG, I extend collective thanks to everyone who travelled to the summit, made presentations, facilitated discussions, provided feedback and otherwise contributed to the success of @Risk North 3.

Through this report, we wish to share and continue the conversations that took place at the summit. Despite our best efforts, we ask for your understanding regarding any errors or omissions in the notetaking. We hope you find the report useful.

Geoff Harder (University of Alberta)
Chair, Digital Preservation Working Group

Appendix 1. Lightning Talks

View the Lightning Talk Presentation Slides as provided by presenters.

Natalie Vielfaure (Manitoba), Making Digital Preservation More [Environmentally] Sustainable.

This lightning talk explored how the integration of environmentally sustainable practices in digital preservation work can lessen the burden on digital preservation practitioners and make financial and human resources more sustainable. Examples on how the University of Manitoba Libraries have started to integrate sustainable practices to reduce their backlog, engage record creators, encourage broader institutional changes, provide new opportunities for partnerships, and reduce the digital preservation workload were discussed.

Julie Shi (Scholars Portal) and Leanne Olson (Western), Scholaris for the Long Haul: Thinking Through Preservation for a National Institutional Repository Service

Both institutions and service providers encounter challenges to preserving IR content, including developing policies and procedures, building technical workflows to enable those procedures, considering the varying needs of diverse outputs, managing associated costs, and more. Scholaris is a new, national, opt-in shared repository service being developed with Canadian Association of Research Libraries (CARL), Ontario Council of University Libraries (OCUL), and the University of Toronto Libraries (UTL), in collaboration with regional consortia and the broader repository community to support open discovery, management, sharing, and preservation of Canadian scholarship. This community-informed approach includes expert groups that are being hosted in partnership with CARL. One such group is the Scholaris Digital Preservation Expert Group (S-DPEG), which was formed in August to provide guidance, recommendations, and advice on digital preservation planning requirements and pathways for the service. This talk outlined the group's goals and provided details on work to date.

Amanda Tomé (Digital Research Alliance of Canada), Fun with File Formats in FRDR: File Format Signature Development for Research Data Formats

In March 2024, the Federated Research Data Repository (FRDR) conducted a file metric scan on datasets in its repository. The results of this scan indicated that most file formats in the repository were not available in format registries and therefore could either not be identified by file format identification tools or were misidentified by the same tools. This lightning talk explored the work undertaken by FRDR to create file format signatures for submission into PRONOM to help identify files that

were identified as unknown or files that were misidentified. It highlighted how file format signature development is approached, how this work will help the larger digital preservation and research data communities, and next steps for file format work in FRDR.

Pascale Montmartin (BAnQ), Our Access-Based Information Model (Original title: Notre modèle d'information basé sur l'accès)

Answering your call for proposals for lightning talks for @Risk North 3, I will present a tool set up to efficiently extract files from our preservation repository. This tool, which uses a public URL as a search criterion, is based on our access-based information model, which will be briefly introduced at the start of the talk. For more details on our information model, here is a link to the paper presented in September 2024 at the iPres conference: https://ipres2024.pubpub.org/pub/mhrmk885/release/1.

Nicholas Lobraico (Canadian Heritage Information Network [CHIN]), Lessons Learned Supporting Digital Preservation in Canadian museums

The Canadian Heritage Information Network (CHIN) is a special operating agency sharing a strategic plan with the Canadian Conservation Institute (CCI). CHIN manages Artefacts Canada, a national database of museum objects; Nomenclature, a standard for museum cataloguing; and specializes in supporting Canadian museums in their digital transformation. Since 2014, this has included publishing resources and, since 2017, offering workshops and webinars on digital preservation for museums. This presentation addressed CHIN's work in supporting digital preservation in Canadian museums, highlighted some lessons learned along the way, and identified next steps to improving services in the future.

Lisa Lawlis (Archives and Special Collections, Western Libraries), Lessons Learned from Implementing Best Practices for Digital Forensic Workflows

This lightning talk focused on the set up (hardware and software) for a processing workstation for archivists at a medium sized university archives and a high-level demo of the forensic workflows for processing and ingesting born digital records. One of the most difficult parts of digital preservation is starting. This presentation shared lessons learned implementing best practices for digital forensic workflows and moving from "theory" to "practice". The talk highlighted a resource guide the speaker is creating to help small to medium sized archives to jump into digital forensic workflows to preserve the Canadian born digital record. The resource guide (title pending) will help other institutions move past "Introduction to Digital Preservation" into practice.

Kenton Good (Digital Production and Preservation, University of Alberta Library), Using the Internet Archive Platform to Rescue at Risk Legacy Digitization Projects

This talk discussed recent work to rescue an offline legacy digitization project (the Kinji Imanishi Archive) by using the Internet Archive platform as a conduit to re-establish both access and preservation for this material. The speaker also touched on conducting such work at scale by using the recent example of the University of Alberta Library migrating 70,000 digitized historical newspaper issues to the Internet Archive.

Robyn Stobbs (Athabasca), Developing Support for Research Data Preservation: Collaborations and Early Steps

Athabasca University has been considering how to support research data preservation and the scope of service offered. This process began with discussions stemming from the Research Data Management Strategy used to build cross unit collaborations to support the development of a research data classification and guidance, and to begin discussions of how research data can (or should) be treated within records retention schedules. The presentation covered the draft research data classification and key points of discussion for different units within the university as well as how this work is being used to identify needs more broadly for digital preservation work, and the infrastructure to support it, at the institution.

Art Rhyno (Windsor), Large Scale Newspaper Digitization with a Teeny-Tiny Budget

OurDigitalWorld (ODW) inherited a community newspaper initiative from a provincially funded, multi-sector project in 2013. As a standalone non-profit in the Canadian cultural heritage sector, ODW has needed to embrace economical solutions for providing digitization services. At the same time, there has been a huge demand for supporting newspaper digitization projects. This lightning talk describes one approach for digitizing newspaper collections by leveraging some of the most affordable options available to Canadian organizations, including low-cost web storage, container computing, vector indexing, and serverless pathways through the rich IIIF ecosystem.

Tanis Franco (Toronto Metropolitan University) and Jessica Ye (University of Saskatchewan; previous Digital Archivist at TMU), Implementing a Digital Preservation Program at Toronto Metropolitan University

The presentation consisted of a case study on launching the digital preservation program at Toronto Metropolitan University, Archives and Special Collections, including the implementation of Permafrost, a digital preservation service operated by Scholars Portal. The presentation focused on the early challenges and key decisions

that shaped its success. The program, begun in February 2024, aimed to safeguard valuable academic and institutional records, addressing the growing need for long-term access to digital assets. The talk covered initial assessments of existing collections, infrastructure needs, stakeholder engagement, and policy development. Emphasizing practical insights, it highlighted how collaboration between IT, library staff, and faculty was pivotal in overcoming resource limitations, establishing preservation workflows, and ensuring sustainability in a rapidly evolving digital landscape.

Julie Shi (University of Toronto/Scholars Portal) and Meghan Goodchild (Queens/Scholars Portal) The Data's Deposited, What Next?: Building Blocks to Support Institutions with Research Data Preservation

Academic libraries are increasingly recognized as partners in and providers of research data management services. In parallel, researchers are increasingly required by funding bodies and publishing venues to deposit and share their data in order to support transparency, reusability, and reproducibility. Underpinning these trends are the FAIR principles of findability, accessibility, interoperability, and reusability, goals that require adequate technical infrastructure as well as ongoing and informed interventions to ensure that research data remains discoverable, accessible, usable, and understandable into the future. Preserving research data to support these principles reveals a number of challenges, including developing policies and procedures, defining workflows and the roles that various stakeholders play in these processes, considering needs across the heterogeneous and evolving landscape of research data (e.g. big data, sensitive data, disciplinary practices and protocols), as well as managing the costs associated with curation, repository management, storage, ongoing preservation, and more.

Presenters discussed efforts to build blocks in support of research data preservation in Canada. These building blocks are of two varieties: (1) ongoing activities at Borealis to technically support institutional preservation priorities, and (2) a project-in-progress with the Alliance's Preservation Expert Group to develop documentation and resource in support of institutions interested in benchmarking or certifying their services based on the CoreTrustSeal certification.

Sarah Lake (Concordia), Preserving Open Research and Unique Archival Materials: Archivematica Implementation at Concordia University Library

This talk provided an overview of how Concordia Library’s Archivematica implementation has evolved in the 5 years since it was launched. Significant milestones include the development of an open-source integration with EPrints, a migration to open-source cloud storage through the Ontario Library Research Cloud,

and the development of efficient processing workflows to successfully tackle large preservation backlogs.

Elizabeth Shaffer (School of Information, UBC), Centering Sustainability in Digital Preservation Education

The ubiquity of rapidly evolving digital technologies at a time of increased calls for climate justice demands that digital preservation education go beyond questions of access and storage to address the critical issue of sustainability. As archives increasingly move to digital formats, the long-term viability of records depends not only on technological and policy infrastructure but also on ecologically and ethically sustainable practices. This talk described developing pedagogy to centre sustainability in digital preservation education, touching on emphasizing the environmental impact of digital preservation, the role of resource management, and socially responsible archival/preservation practices by integrating sustainability into digital preservation curricula.

Curtis Frederick (University of Calgary), Preserving the Digital Footprint: Navigating Web Archiving Challenges in CMSs at the University of Calgary

Many institutions rely on web content management systems (CMS) to share knowledge with the public. This information often holds archival significance and is a prime candidate for long-term digital preservation, even if it resides on the CMS for only a short time. While some content may exist in various formats or systems, much of it is created and edited solely within the CMS, complicating the archiving process. This presentation looked at two cases studies at the University of Calgary that have attempted to use web archiving as a solution for archiving and preserving records and information from CMSs. The University of Calgary has recently explored and experimented with using web archiving to capture embedded PDFs of course outlines and capturing news articles written directly in a CMS. The presentation looked at our successes and failures, examined the limitations of using web archiving, and highlighted the importance of archivists proactively collaborating with individuals and departments from the outset to ensure that these valuable records are effectively captured, archived, and preserved.

Paul Durand (Military History Research Centre, Canadian War Museum), Digital Heirlooms – Canadian War Museum

The Canadian War Museum's National Collection is largely collected through public offers of donation. The museum receives over 600 donation offers a year, representing 1000's of objects. Each year the museum's archive accepts 100's, and sometimes 1000's, of personal letters, diaries, photos, scrapbooks and films that tell the story of Canadians affected by conflict. They are often objects put aside by an

individual and passed down as family heirlooms. They may be proudly displayed on a bookshelf, or more often than not, hidden in a closet for decades until rediscovered. Family heirlooms are often offered to the museum decades, or even 100 years, after their creation.

Andrea Mills (Internet Archive Canada), Internet Archive Updates.

N/A

Appendix 2. Digital Preservation Tools and Technologies – Insights

Digital Preservation Responsibility	Solution or Approach	Strategies and Insights
Data Integrity and Authenticity	Tools like DROID, JHOVE, and checksum utilities (CRC, MD5) are valuable for ensuring data integrity during transfers.	Challenges exist in validating backups; test restores are critical to ensure reliability. Multi-layered backup strategies with geographic separation and regular testing are emphasized.
File Format Management	Strategies to manage obsolescence include tools like DROID, Siegfried, and Local Digital Format Registries.	Guidelines to steer donors toward open and stable formats can prevent future access issues. Could enhance workflows to deal with difficult formats before they become obsolete.
Scalability in Storage	Could benefit from adopting scalable tape storage solutions like LTO 9, which are cost-effective for high-volume data.	Advocacy for necessary storage resources is essential, as IT may underestimate archival storage needs.
Migration of Analog Materials	Proven tools for digitizing analog materials include ABBYY FineReader, Adobe Bridge, and in-house scanning solutions.	Documentation and knowledge-sharing are critical to address challenges with single knowledge keepers.
Complex Digital Assets	Tools like Archive-It, ePADD, and Aid4Mail support managing websites, social media, and emails.	Normalizing databases to standardized formats like CSV enhances long-term usability.
Metadata Management	Standardization using schemas like PREMIS and tools	Maintaining a master metadata copy (e.g., spreadsheets)

	like Archivematica, OpenRefine, and Python scripts ensures consistent metadata capture.	simplifies management and long-term preservation.
Automation and AI	AI tools like ChatGPT, Whisper, and Python scripts streamline metadata generation and repetitive tasks.	Can explore AI for transcription, accessibility, and anomaly detection in its digital archives.
Collaboration and Tools Sharing	Leveraging collaborative platforms like Permafrost and Scholars Portal can enhance coordination with other institutions.	Open-source tools and community-maintained repositories are increasingly important for scalability and access.
Training and Adoption	User-friendly tools like Audacity and Preservica are easier to adopt, but staff training remains critical.	Developing in-house Python expertise can significantly improve automation and tool adoption.

Appendix 3. Key Challenges and Opportunities in Digital Preservation for the Next Five Years

Area of Concern	Challenge	Opportunity
Technology, Processes, Preservation	Internet, social media, and AV content are challenging Obsolescence Workflow Technological barriers Ethics for harvesting and for access Evaluating authenticity of content for preservation given prevalence of AI-generated information Preservation of AI models Environmental footprint of cloud storage Address capacity and high-volume entering repositories Technical scalability and integration	Share and adapt methodologies; create case studies Roving digital preservation specialist to offer audits Assessment model created by CARL, adapted from DPC ram assessment model Encourage more inter-institutional collaboration in area of cybersecurity Do a gap analysis, compare institutions for benchmarking Collect and share best practices, policies and strategies Use AI to enhance metadata, identification and appraisal Train more reliable AI using GLAM sector content Create guidelines for capacity
Sustainability	Flat or decreasing budgets Need Canadian solutions, not just local ones, to maintain sustainability Precarity of grant-funded projects	Community funding opportunities, such as SSHRC communities grant or funding from Desjardins

Advocacy	Grow understanding of the importance of access in the present and also for the future Government and institutional relations Advocate for digital preservation with limited time and resources	Establish a funded position to advance digital preservation nationally Establish a national oversight body for digital preservation Policy leadership from larger institutions like LAC and BAnQ Change language from “digital preservation” to “research information security” Create communication tools to raise awareness with non-experts, i.e. administrators, other employees, members of the public
People	Better define responsibilities Need for indicators of success	Establish a community of practice; could also liaise with DPC or CRKN Establish a network of mentors/mentees Foster inter-institutional mentorship between large and small institutions Create a national network or list of experts Increase frequency of digital preservation summits Offer summits in conjunction with training opportunities Create training competencies Implement training competencies beyond libraries and archives

		Increase training in college and university programs Share job descriptions for digital archivists and “informaticien de préservation”