

## Initiatives

Outcomes from the *Building a Distributed Digital Library* session on Nov 1, 2018 for consideration

### **Digitization**

The BC Digital Library will undertake a multi-collection digitization initiative as a two year proof of concept commencing in early 2019. It is envisioned as the first step in realizing the BCDL's vision to enrich the intellectual and leisure life of a citizen, showcase the contributions of individuals, organizations and government, and contribute to the economic well-being of the province. The digitization initiative will be a collaborative undertaking that matches up the existing expertise and digitization resources available at some of the larger sites -- RMBC, SFU, UBC, UVIC -- with BCDL participants and community partners who are interested in digitizing and making accessible collections that are well suited as first-time digitization projects. The intent is to identify at least 15-20 suitable projects that would be submitted to the BC Historical Digitization Program for funding.

BCDL proposes to solicit a diverse range of potential participants with appropriate collection candidates for the BCHDP. The focus would be on those sites with limited, if any, digitization experience and insufficient technology resources to undertake this work independently. Each participant would be matched up with one of the larger participating sites and the two would jointly prepare a proposal for submission to the BCHDP. The first-time sites would bring a unique collection along with their commitment to be immersed in a digitization project and gain valuable experience. The senior site would provide consulting support on preparing the BCHDP proposal, permission/copyright requirements, and best practices for digitization projects. They would also provide the technological resources and infrastructure for digitizing, OCRing, loading, and making the content accessible online via their infrastructure, through shared services such as ARCA, and/or another local platform.

All parties will derive benefits from these collaborative digitization projects. The first-time sites will obtain invaluable experience in all facets of a digitization project that would produce a tangible outcome -- a fully digitized, openly available collection available online, including a long term preservation solution for their digitized content. The senior sites will be able to share their expertise and technical infrastructure on a cost-recovered basis via the BCHDP grant. The BCHDP will have 15-20 high-quality proposals that will follow digitization best practices with a guaranteed and persistent outcome. Everyone will reap the benefits of 15-20 new, unique BC-based digital collections freely and readily available online for research, educational and community purposes.

For BCDL this initiative will be an opportunity to advance the vision to build a comprehensive set of BC-based digital resources freely available online for everyone, but especially for BC residents. It will be a major "proof of concept" step that will explore potential service offers, strengthen relationships among stakeholders, and provide a pragmatic basis for further consideration of the most suitable governance, operational and business models to realize the

full potential of this undertaking. It will be solidly grounded in the extensive practical experience and resources the senior sites will bring to the table. In many respects, it will simply scale-up the number and variety of collaborative digitization projects that have already been taking place amongst a variety of the current and potential stakeholders irrespective of the size or type of organization.

### **Crowdsourcing**

As the BC Digital Library (BCDL) initiative grows its founding members will plan to add interactive, crowdsourcing mechanisms to the public interface. Internally to its members, the BCDL is a service offering provided by cultural institutions for cultural institutions. The services focus on cultural heritage digitization and preservation as a logical starting point, but they will not end there. Externally, the most important service will be ongoing public access.

Public access can take many forms. Traditionally, digital heritage is pushed out to the public in form of digital imagery, interpretive text and metadata. The push is one way. Insight that could be gathered from an individual who experiences digital culture is not captured, and cannot enrich the original record. If we consider the digitization and cataloguing backlogs faced by the majority of Canadian memory institutions, a one-way push is a two-fold loss. Cultural records are not improved, and citizens do not have an opportunity interact more directly with digital artifacts. In 2015, an expert panel on memory institutions and the digital revolution concluded, “memory institutions can leverage both skilled and non-skilled input from citizens while providing them with enhanced and valued experience.”<sup>1</sup> The success of the BCDL should be measured by its value to local citizens and wider digital communities. To ensure that success the initiative will go beyond traditional one-way communication with its audiences. The BCDL will go beyond even the notion of open, free access. Free access implies *take this and do with it what you will*. It does not provide an incentive for citizens to create lasting relationships with memory institutions and their digital cultural heritage. To form trusted relationships with the public the BCDL should plan to build mechanisms for discourse related to digital artifacts directly into the public platform.

### *Two Approaches*

Initially, a more self-regulated and automated crowdsourcing approach is proposed. Crowdsourcing cultural content can range from free-form, personal storytelling to a less individualized experience with original source material like transcription or proofreading. The goal of the latter is make corrections or an index of the records. One example is to allow the public to transcribe cursive letters into typed text. The benefits are enormous. Proper names, cities and towns, and dates are all now indexed for searching. The transcriber has the satisfaction of knowing the contents intimately, and improving access to the material by researchers and historians. Another highly successful example is from Trove - Australia’s national library and digital heritage portal. The majority of their collections are scanned newspapers that have gone through an optical character recognition (OCR) process. OCR is mildly accurate at its

---

<sup>1</sup> Owran, Doug, et al. “Leading in the Digital World: Opportunities for Canada’s Memory Institutions.” Council of Canadian Academies. (2015). Retrieved from: [https://scienceadvice.ca/wp-content/uploads/2018/10/cofca\\_14-377\\_memoryinstitutions\\_web\\_e.pdf](https://scienceadvice.ca/wp-content/uploads/2018/10/cofca_14-377_memoryinstitutions_web_e.pdf)

best, and so Trove allows citizens to proofread and perform typo edits. The uptake by citizens is impressive. A dozen or more users have corrected over 1.5 million typos each.<sup>2</sup>

Eventually the BCDL should look to other national and international examples of crowdsourcing that allow for more individualized experiences with cultural objects. Europeana and DigitalNZ both allow end users to collect desired objects and interpret them based on personal experience.

### *Cart Before the Horse*

The potential for crowdsourced, cultural content is unlimited. Before it's realized the BCDL initiative must make the transition from talk to action, from discussions about services to providing those services, and establishing a streamlined executive and administration team to monitor operations. Crowdsourcing, even at a small level, requires dedicated resources to moderate new content for relevance and decorum. The staffing levels needed are not onerous or out of reach the BCDL's GLAM collaborators, and the benefits vastly outweigh the effort.

### **Linked Open Data**

The vision of BCDL is to bring together collections from the GLAM sector so that they might be flexibly represented through a discovery interface. Applying linked open data in this environment is an opportunity to enhance basic search with semantic search capability, facilitating discovery of individual collections but also making connections among collections through shared and related meaning. Linked open data allows integration of our collections at the content level and de-silos collections that currently exist side-by-side but still separate. As a regional cultural heritage hub, BCDL is an ideal context to apply linked open data technology and semantic logic.

Scaling linked open data is the challenge. BCDL could explore ways forward through a contributing project of at least two participating institutions that have the knowledge and resources to apply RDA metadata to records. For example the Official Gazettes of the Province of British Columbia from 1863 to 1871 at University of Victoria and the BC Historical Newspapers collection from UBC would be a powerful proof of concept and capacity exercise for BCDL, and could be considered in relation to means of creating thematic collections or sets that are used by other organizations with similar mandate (e.g Europeana, DPLA).

Such a project is not prerequisite to other technical work related to development of the BCDL platform, but may be undertaken to enrich it, through the in kind contribution of willing participants with expertise. It would require regular communication with BCDL admin and developers about technical matters and the progress of the project, points of integration with the platform in development, communication of findings.

---

<sup>2</sup> Trove: National Library of Australia. (2019). Retrieved from: <https://trove.nla.gov.au/newspaper/hallOfFame>