



# Library and Archives Canada

## Rethinking the Stewardship of Newspapers in a Digital Age

### 1. Introduction

This document should be read in the context of the three following documents, which are available on the Library and Archives Canada website ([link](#)) Library and Archives Canada's Story Continues; the Documentary Heritage Management Framework; and the Acquisition Orientation Instrument.

This document is part of a series of eight discussion papers that seek, through practical projects, to validate the four guiding principles and the key roles of the Documentary Heritage Management Framework developed by Library and Archives Canada (LAC) during summer and fall 2009. Specifically, each pathfinder project seeks to demonstrate the applicability of these principles and roles in the context of the acquisition business pillar. Each pathfinder project will result in a final report, which will add overall conclusions about the implementation of the Documentary Heritage Management Framework and the Acquisition Orientation Instrument. Based on observations, the final report will be even more effective in contributing to the achievement of LAC's mandate, within the current operating environment.

### 2. Background

The newspaper industry has changed significantly in recent years and continues to evolve at a rapid pace. Ownership has consolidated; production has turned digital; online dissemination has taken hold; traditional formats and business models are threatened; citizen journalism is challenging the role of professional journalists; and content is increasingly repurposed, fluid, and interactive.

Memory institutions have traditionally collected newspapers for the benefit of their user communities, providing free access to and preservation of this massive body of contemporary and historical content. The stewardship of newspapers, however, is an increasingly complicated and challenging task requiring memory institutions to rethink how they select, collect, preserve, and provide access to newspaper content. As the newspaper industry continues its turbulent transition to the digital environment, memory institutions are, with limited resources, transitioning their own newspaper stewardship practices.

To respond to this changing environment through modernization, Library and Archives Canada (LAC) initiated a pathfinder project to examine the stewardship of the national newspaper collection. Stewardship refers to the processes of selecting, collecting, managing, preserving, and providing enduring access to this body of Canada's documentary heritage within a framework of respect for intellectual property rights. This pathfinder project has the following objectives:



1. To examine and make recommendations related to Canadian newspaper<sup>1</sup> acquisitions, collection management, digitization, and long-term retention in consideration of LAC's recently articulated four guiding principles of significance, sufficiency, sustainability, and society.
2. To study and define a feasible path toward a digital model for access and preservation of Canadian news media content, in consultation and cooperation with key stakeholders within the context of the shift to digital production and online use.

### 3. Current State of Newspaper Stewardship

As Canada's foremost memory institution for preserving the nation's documentary heritage, Library and Archives Canada (LAC) holds the most comprehensive collection of Canadian newspapers in the world. While LAC has never attempted to comprehensively collect and preserve Canadian newspapers, its newspaper collection is broad and extensive, comprising:

- An extensive collection of Canadian daily, weekly and community newspapers on microform;
- Complete retrospective runs of 10 Canadian major daily newspapers in print form;
- First, last and special editions of major newspapers in print form;
- Collections of student, Aboriginal and multicultural newspapers in English, French, and many other languages in print form.

Until now, LAC has not been actively collecting online newspapers, nor has it had a program to digitize its newspaper collection. However, LAC is now well advanced in its efforts to become a Trusted Digital Repository (TDR) capable of ingesting, managing, and preserving digital content into the future; as such, it is looking to increase its digital newspaper offerings.

Newspaper content currently exists in multiple formats. The following sections outline the current state of print, microfilm, digitized, and born-digital newspaper stewardship, as well as the types of access available for newspapers.

#### 3.1. Print Newspapers

Owing to their volume, print newspapers are not subject to the provisions of legal deposit under the *Library and Archives of Canada Act*. However, LAC has maintained subscriptions to a broad range of Canadian newspapers, some of which are permanently retained in print as part of Canada's documentary heritage collection.

LAC maintains subscriptions to 10 Canadian dailies which, according to a policy established in 1969, it retains in print form as a representative archival sample of Canadian major newspapers in French and English from different parts of Canada. These titles are currently bundled and stored

---

<sup>1</sup> It should be noted that this initiative is scoped to "print news media" in newsprint, microform, and digital forms; it does not include broadcast media.



in offsite storage facilities. The current subscription cost for these 10 dailies is \$6,781 per year.

In addition to the 10 dailies, LAC subscribes to 171 Canadian print newspapers at a cost of \$32,740 per year. These subscriptions include: 107 multicultural newspapers, 17 Aboriginal newspapers, and 47 other titles representing various cities and communities from across Canada. The print versions of the 47 representative titles are discarded once the microfilm versions are received. LAC also collects a selection of non-subscription university student, multicultural, and community newspapers.

Currently, parts of this massive print collection are stored at the Library and Archives Canada building at 395 Wellington Street, as well as at three other storage facilities in Ottawa. To address collections at risk, such as the fragile print newspaper collection, and to improve the storage conditions of these collections, LAC has received one-time special funding from the Government of Canada to convert an industrial building into a collection facility, and this will house the LAC newspaper collection. The print newspaper collection will be moved at the beginning of 2012. The move of this collection will cost an estimated \$500,000 owing to the collection's scale, fragility, and the need for substantial re-boxing. There are an estimated 80,000 bundles of fragile, original newspapers to be boxed, with two bundles per box. There are also about 6,000 oversized bound volumes which, while in generally good condition, are also to be boxed.

### **3.2. Microform Newspapers**

All Canadian newspapers published in microform after 1988 are subject to legal deposit. Two copies are sent to LAC when more than 100 copies are produced, and one copy when the run is between 4 and 100 copies. LAC purchases titles filmed for preservation purposes (fewer than four copies produced) as funds permit, as well as titles filmed before 1988. In fact, LAC spends more than \$90,000 per year on newspapers on microform.

The number of Canadian newspaper titles published in a year and over time does not make it feasible for one institution to collect every newspaper. As a result, the Decentralized Program for Canadian Newspapers was established in 1985. Under this program, which ceased in the late 1990s, the provinces and territories were responsible for acquiring, filming, and listing their respective newspaper titles allowing for the disposal of print originals. LAC was responsible for compiling a national list (the Union List of Canadian Newspapers, now included in the AMICUS database), as well as facilitating national and international lending, location, and reference services.

This successful program built strong provincial newspaper collections within one or more memory institutions in each province and territory. LAC effectively subsidized the filming costs by committing to purchase a copy of all newspapers on microfilm produced, and through this means, built an unmatched pan-Canadian microform newspaper collection for all Canadians.

While subscription costs for newspapers in print and microform total over \$129,000 per year, the total cost of ownership including storage, preservation management, and description is much higher. Given this, questions arise as to whether the cost of print storage is still justifiable or if there are other alternatives? Furthermore, when is it justifiable for LAC to continue print subscriptions?



### 3.3. Digitized Newspapers

In Canada, projects to digitize historical newspapers began to emerge as memory institutions and some private companies recognized the access opportunities of the Web. Since the late 1990s, there have been numerous retrospective newspaper digitization projects in Canada, of which some major ones include: Paper of Record, Pages of the Past, The Alberta Heritage Digitization Project, Multicultural Canada and the digitization work of Bibliothèque et Archives nationales du Québec (BAnQ).

The following is a summary of observations of the state of Canadian newspaper digitization based on LAC's own newspaper digitization experience as well as an examination of select major initiatives in Canada.

- While LAC has digitized some special editions of Aboriginal and French-language newspapers, as well as Canada's first newspaper, the *Halifax Gazette* and the *Canadian Illustrated News*, LAC has not pursued a program of mass digitization of newspapers to date. Available funding levels have supported only exhibition-scale digitization, not the digitization of full back runs of numerous titles.
- Among Canadian projects, geographic coverage is uneven with the most digitization taking place in Alberta and Quebec.
- Access models vary based on copyright and business models. Most copyright-cleared historical content that has been digitized is freely available, but there are a few business models like *The Toronto Star* and the *Winnipeg Free Press* that charge for access to archived content.
- LAC has not attempted to acquire on legal deposit Canadian newspapers that have been digitized by other organizations. Depending on the digital capacity of the host institution, long-term access to some of this content could be at risk.
- When the Cold North Wind Corporation's extensive collection of digitized newspapers was made available for sale, the new legal deposit regulations were not in place and LAC was not able to afford its purchase price. It was subsequently sold to Google. As of June 2009, it is reported that Google has made 522 titles (4.9 million articles)<sup>2</sup> freely available.

The level of Canadian activity and investment to date does not compare favourably with the newspaper digitization of other countries. International efforts reveal certain trends, emerging best practices, and some key knowledge about the methods, technologies, costs, and comparative merits of various approaches to newspaper digitization. Key observations of these international initiatives are:

1. There is a clear shift from boutique to large-scale digitization.
2. Funding comes mainly from governments and grant programs.
3. Most organizations are digitizing out-of-copyright, historical newspapers.
4. Most organizations are outsourcing newspaper digitization, and are digitizing from microfilm whenever possible.
5. The scale of output ranges significantly, but in each case it is substantial ranging from one to eight million pages.

---

<sup>2</sup> The Cold North Wind titles are included in the 522 newspaper titles in the Google News Archive. While not all of the Cold North Wind material was Canadian, there was a large amount of Canadian coverage, most of which is now available via Google News Archive.



6. There are significant differences in costs depending on the type of post processing undertaken (i.e., whether full-text search is offered or only browsing access).
7. There is a mix of collaboration models including private, government, non-profit, and university partners.

Given the current state of newspaper digitization, what actions should LAC and other memory institutions take to advance the digitization and access to a coherent, pan-Canadian retrospective collection of Canadian newspapers?

### 3.4. Born-digital newspapers

While online newspapers have been subject to legal deposit since January 2007, LAC has yet to begin actively collecting these websites. This is primarily due to the fact that until now LAC has not had the systems required to ingest, store, manage, preserve, and provide access to such voluminous, dynamic, and rights-encumbered content. However, LAC is now well advanced in the development of its Trusted Digital Repository (TDR), into which digital content can be ingested and thereafter managed and preserved. The functionality to ingest serial publications to LAC's TDR is expected to become operational by the summer of 2010. However, while considered a form of serial publication, online newspapers present special complications.

- First, their constantly updated news cycle means the concept of an 'edition' no longer applies, raising the possibility that even a daily capture of a newspaper's website would likely miss some content.
- Second, online newspapers increasingly enable Web 2.0 forms of citizen journalism and social commentary from the readership – content that is also highly fluid and may be less appropriate for long-term capture and retention.
- Third, newspaper content is by no means unique on a title-by-title basis. The consolidation of news media ownership has also led to content being supplied to a company's chain of papers from a single corporate-level feed, leading to considerable overlap and redundancy among the content of daily newspapers published across the country. This is most clearly illustrated by CanWest and the way in which its central site presents its 12 major Canadian dailies. In addition, newswire services such as CP, AP and Reuters may furnish up to half of the news content of a given paper, adding another layer of complexity and possible content redundancy to the picture.

These factors combine to challenge LAC's selection decisions for long-term retention. Owing to the wealth and breadth of its contemporary documentary content, newspapers—and in particular major dailies—are recognized to be a significant form of documentary heritage. But, as with print, not all online newspapers are equal in their degree of significance. What coverage across the wide range of possibilities is a sufficient capture of the newspaper content that Canadian society is generating? What amount of digital newspaper capture will be sustainable by LAC, now and into a long-term future?

For efficient capture, harvesting technologies offer some hope, and LAC has built some technical capacity and experience in this regard by conducting a variety of other forms of Web crawling (Government of Canada web domains, elections sites, Olympics sites, etc.).

### 3.5. Copyright and Newspapers

Newspapers present a particularly complex set of rights challenges, as the rights holders can include newspaper owners, freelance contributors, and photographers. For articles written by



staff writers, the copyright typically remains with the newspaper. Most newspapers are written by a combination of staff writers and freelance contributors or the articles are received through newswire services. While it depends on the terms and conditions of each contract, freelance writers can hold the rights to their own content. In general, the term for copyright is the life of the rights holder plus 50 years.

In addition, there are also some unique digital rights issues. In 2006 the Supreme Court of Canada made a decision regarding the use of newspaper content in databases that may have broader implications for digitization. According to the *Robertson v. Thomson Corp.* case, newspapers do retain the right to reproduce content including freelance content, provided that they do not de-contextualize the articles to the point that they are no longer “presented in a manner that maintains their intimate connection with the rest of the newspaper.” This ruling indicated that while newspapers had the right to reproduce something that looked like the original newspaper, they did not have the right to reproduce individual articles in a database without the permission of the author.

In light of this complex rights scenario, the de facto guideline for considering newspapers copyright clear in Canada has been 90 years, and so, memory institutions have tended to concentrate on digitizing newspaper content published before 1920.

In terms of collecting online newspapers, the legal deposit regulations do give LAC the right to collect online newspapers for preservation purposes. Permissions for access would have to be negotiated with the rights holders.

Whether collecting current online newspapers or digitizing retrospective newspaper content, a future stewardship model must respect the rights of all parties involved, and should seek to balance the interests of the newspapers, the content creators, and the public.

### 3.6. Access to newspapers

The LAC newspaper collection in microform format is available to Canadian and non-Canadian researchers for onsite consultation and through interlibrary loan. This collection constitutes one of the most heavily used parts of the LAC collection. LAC maintains on its website a list of its holdings of Canadian newspapers in microform format, organized by province or territory and by city. This list contains over 2,300 Canadian titles including multicultural, Aboriginal, and student newspapers. All newspapers are catalogued and included in the AMICUS database. Print newspapers can be consulted in the reading rooms at 395 Wellington Street in Ottawa; however certain restrictions may apply depending on the condition of the newspaper and the purpose of the request. Most of the requests for print newspapers are for the first, last, and special editions.

As the new Web technologies continue to develop and more newspaper content goes digital, we are seeing an increasing array of access options. Contemporary access models for digital newspapers include:

- Access to current news coverage via a newspaper’s website. Depending on the business model, this type of access can be either free or subscription based.
- Access to digitized historical content via a newspaper’s website. While the majority of online newspaper sites tend to only offer access to contemporary content (approx. 1990s onwards), some newspapers have digitized their content and offer subscription access to their archives. Examples include the *Winnipeg Free Press* and *The Toronto Star*.



- Subscription access to a news database. While most major Canadian daily newspapers are available through database services, the focus of their archives tends to be quite contemporary (approx. 1980 onwards) and access is by subscription. Additionally, articles in the subscription services tend to be text only and do not have images, ads, classified notices or the original layout that may have social and historical research value. Some databases, however, do offer full-text access to historical content as well. For instance, ProQuest Canadian Newsstand, offer full-text access to 240 Canadian newspapers covering from about 1977 to present.
- News search aggregator. In recent years, services such as Google News have become prominent. Such services crawl and index online newspaper content allowing users to keyword search contemporary news content. While the index allows users to search across a vast amount of news content, access is a mix of free and subscription. In the case of Google News, most of the content is held on the newspaper's website and just indexed by Google; however, Google has started to purchase and provide access to historical content. In 2008 they purchased over 700 digitized newspaper titles from Cold North Wind and are currently integrating full-text access to this content via Google News Archive, which is currently a free service.
- Thematic aggregation of digitized content. Some initiatives such as Multicultural Canada have digitized historical as well as recent content and are providing free full-text access via a thematic website.
- Access to digitized regional content. Digitized historical content is also aggregated into regional collections. Peel's Prairie Provinces brings together bibliographic data as well as full-text content dedicated to the exploration of western Canadian history.

While access is increasing through new sites and services, preservation concerns remain. Similarly, content available in subscription databases is limited in coverage (not a full run), and while access can be licensed there is no guarantee of long-term preservation and access. In short, while news aggregator services like Google News and subscription databases such as ProQuest and Gale provide easy access, they do not have a requirement to address the preservation requirements of long-term stewardship.

## 4. Going Forward: Toward a New Stewardship Model

Two primary factors are driving the need to reassess newspaper stewardship in this country. First, newspaper publishing has turned to digital form. Production is entirely digital; while distribution is increasingly, but not exclusively, digital – at least not yet. This valuable “born-digital” content will be lost forever if a selection of it is not stewarded now and into the future by committed organizations that have the requisite organizational, technical, and financial capacity to be a Trusted Digital Repository.<sup>3</sup>

The second factor is that use is now predominately online. Researchers and the general public turn to the Internet to find the information they want, conduct their research, and discover, view and read what interests them.

---

<sup>3</sup> See *Trustworthy Repositories Audit & Certification (Trac): Criteria and Checklist*.  
<http://catalog.crl.edu/search~S1?/.b2212602/.b2212602/1,1,1,B/1856~b2212602&FF=&1,0,,1,0>



For memory institutions the shift to digital news production, dissemination and use is having a significant impact on approaches to stewardship. With limited resources to expend on newspapers, yet an undiminished commitment to ensuring responsible stewardship of this segment of Canada's documentary heritage, LAC is faced with some critical choices.

The following table outlines some proposed objectives, strategic directions, and next steps for modernizing LAC's approach to national newspaper stewardship. Before finalizing its decisions, LAC will be seeking feedback from stakeholders on its proposed choices.