

# History for High Schoolers: The Atlantic Canada Virtual Archives<sup>[\*]</sup>

Corey Slumkoski

Department of History and Electronic Text Centre, University of New Brunswick

[corey.slumkoski@unb.ca](mailto:corey.slumkoski@unb.ca)

CHWP C.4, publ. January 2007. © Editors of CHWP 2007.

---

**KEYWORDS / MOTS-CLÉS:** Digital imaging, pedagogy, technology, text transcription / *Imagerie numérique, pédagogie, technologie, transcription de textes*

---

- [Introduction](#)
  - [1. The Edward Winslow Family Papers](#)
  - [2. The McQueen Family Papers](#)
  - [3. The ACVA Website](#)
    - ◊ [3.1 Developing ACVA: Digital Imaging](#)
    - ◊ [3.2 Developing ACVA: Text Transcription, Markup, and Manipulation](#)
    - ◊ [3.3 Developing ACVA: Learning Resources](#)
  - [Conclusion](#)
  - [Notes](#)
- 

## Introduction

In 2002, Margaret Conrad, the University of New Brunswick's Canada Research Chair in Atlantic Canada Studies (CRC), in cooperation with UNB's Electronic Text Centre, began the development of a portal website designed to provide scholars of Atlantic Canada with "one stop shopping" for resources pertaining to the region. Included in the portal website is a searchable database of works published on the region, an e-print repository, and a wide selection of annotated links on Atlantic Canada. In addition to these features, the CRC also began an ambitious internal program of online publishing which led to the development of the [Atlantic Canada Virtual Archives](#) (ACVA). Funded in part by the Canadian Foundation for Innovations, the ACVA is a digital archive of historical documents. It currently hosts selections from two archival collections: the papers of prominent loyalist Edward Winslow and the McQueen Family Papers.<sup>[1]</sup> This paper will explore the development of the ACVA website, following the project from the idea stage, through the digital imaging and resource development stages, to the finished product.

### 1. The Edward Winslow Family Papers

The Edward Winslow Family Papers contains 38 volumes of correspondence, accounts and ledgers, letter-books, diaries, and notebooks with a chronology ranging from 1695 to 1866. All told, the collection contains over 3000 separate documents in excess of 11,000 pages. The bulk of the collection consists of the correspondences between Edward Winslow Jr. and his compatriots. Winslow, a Mayflower descendant born of a wealthy Boston family in 1745, had lived a privileged pre-Revolutionary life. Educated at Harvard University, he was an entrenched part of the Massachusetts colonial elite. A prominent Boston Tory, he witnessed firsthand the outbreak of the American Revolution at the Battle of Lexington, and was compelled to throw his support to Great Britain in the conflict. Although Winslow journeyed with the British Army to Halifax, where he was appointed Muster-Master General of the American provincial regiments, he spent most of the war in New York, confident of a quick British victory. Winslow's confidence proved misplaced, and with the rebel victory he found himself ostracized from his prior station. The positions of privilege he had long taken for granted would now be denied him. Rather than settle for a position below his perceived station, Winslow, like 75,000 others who had supported the crown during the conflict, chose to migrate to another British holding. In 1782 Winslow was at the vanguard of the over 30,000 Loyalists who traveled to Nova Scotia. As Winslow's close friend Ward Chipman informed him in a July 1783 letter, "Nova Scotia is the rage."<sup>[2]</sup>

Edward Winslow is of particular interest to historians for two reasons. First, he was an integral part of British colonial society. As Muster-Master General during the Revolution he had either seen firsthand or read detailed accounts of the horrors of war. Of even greater interest, however, is the story the papers tell of Winslow's arrival and settlement in British North America. Winslow was a key player in the political and social life of post-Loyalist Nova Scotia. He assisted in the partition of New Brunswick from Nova Scotia, served in a variety of positions in the new government, and was an vital part of New Brunswick high society until his death in 1815 at the age of 70. Winslow's station, therefore, made him privy to a great deal of the comings and goings of colonial society and politics. Second, and no less important, Winslow was a prolific and gifted writer who left behind an incredibly rich series of correspondence. His letters are incredibly engaging and often quite humorous, and through them the eighteenth-century comes alive on the page. Over two centuries later, his papers are still a compelling read.

## **2. The McQueen Family Papers**

The McQueen Family Papers are likewise engaging. Covering the period from 1866 to 1930, they consist of 11,000 pages of material contained in 1200 letters and other documents, such as diaries, account books, artwork, teaching licenses and photographs. The McQueen Family – parents Daniel and Catherine, and children Jane, Mary Bell, Eliza, Dove, Jessie, Annie and George – hailed from Pictou County, Nova Scotia, but they travelled extensively across Canada. The extensive correspondence between the seven McQueen children and their parents, beginning in the 1870s when the oldest children left their rural Pictou County home to study and work, is a fascinating record of life in late-nineteenth and early-twentieth century Canada, and serves as an exceptional window into the Canadian past. The letters that the McQueens wrote to each other are remarkable documents, revealing much about family survival strategies in early industrializing Canada. Five of the six daughters became school teachers, with two of them moving to British Columbia in the late 1880s to seek better paying positions, while their only surviving son moved to New York to seek fortune.<sup>[3]</sup> But the letters do much more than tell of the writers personal lives and

relationships; they also document the story of one family during the years following Canadian Confederation. As such, the McQueen collection is a wonderful resource for the social history of early Canada.

### **3. The ACVA Website**

The [ACVA website](#) currently has two component sections based upon these archival collections. Edward Winslow is featured on one part, the McQueen Family on the other. Both are available in English and French, and adhere to the same basic four-part structure. The first section of each side provides specific background information about the two families. The second section of the site contextualizes the historical experience of the two families, while the third allows users to search and view all the letters held in that particular side of the archives. Finally, the fourth section of each side provides a series of Learning Resources geared toward students in grades 9-12.

#### **3.1 Developing ACVA: Digital Imaging**

The first step in developing the Atlantic Canada Virtual Archives was to digitize the two collections. All digitization was undertaken by the University of New Brunswick's Digital Imaging Centre. Although only the first four volumes of the Edward Winslow Family Papers appear on the ACVA website, all thirty-eight volumes were digitized between 2000-2003 as part of the [Edward Winslow Family Papers website](#). The digitization of the McQueen Family Papers took much less time, and was completed over during autumn 2003. Master archival image files are full colour (24 bit RGB) at a resolution of 300 pixels or dots per inch (ppi/dpi). Following the recommended best practices of leading cultural heritage preservation institutions (for example, [Cornell University's Department of Preservation and Collection Maintenance](#) and [Canadian Heritage's Guidelines for Creating and Managing Digital Content](#)), tonal scale and colour balance controls were set prior to image capture in order both to create digital surrogates that are true to the original documents' appearance and to minimize adjustments during and after processing.

Once images were captured, files were sharpened as needed to achieve the approximate appearance of the original. All sharpening was done in Adobe Photoshop using the unsharp mask filter. Master image files were stored as uncompressed TIFF files (Intel byte order, header version 6). All file-naming for both projects followed established conventions at the University of New Brunswick for effective management of digital image collections, by using a numeric file-naming system based on the source documents physical organization. For example, the Winslow Collection consists of 38 volumes. Each volume contains a number of documents, and each of those has a number of pages. Therefore, filenames were structured volume number > document number > page number. Thus the file-name for the fourth page of the 30th document in the 12th volume would thus be 12\_30\_04 (volume 12, document 30, page 4). By consistently applying this file-naming system one can easily identify exactly where any given file fits in the collection. The same naming conventions, modified to fit the McQueen source documents' physical organization, were used with the McQueen papers.

From these archival TIFF files we created web surrogates for use in on-line delivery. The format for these files is JPEG (24 bit RGB), a flexible, compressed format and the recognized industry standard for the Web presentation of textual and photographic documents. To improve networked access and use of the images

– to speed up the website – resolution of surrogates was reduced to 72 dpi. Additional surrogates in the form of thumbnails were developed for web access. Thumbnails are in JPEG format at a resolution of 72 dpi with reduced dimensions of 150 pixels in width for landscape images and 150 pixels in height for portrait images. The university has ensured that image files can be identified with a persistent URL to enable reliable citation, cross-linking, and integrated access.

Master images (TIFFs) are archived to CD-R while surrogates (JPEGs) have been uploaded to a Unix (Linux) server running Apache Web server software. The Winslow Family Papers had two archival CD-Rs burned – one to remain in the Digital Imaging Centre and one to be given to the University Archives. For the McQueen Family Papers three CD-Rs were burned – one for the Nova Scotia Archives and Records Management and two for the University of New Brunswick.

In order to manage the images in the collections, cataloguing and metadata descriptions were created at collection, document and component image levels according to the Electronic Text Centre's extended Dublin Core metadata schema. The project follows a Dublin Core framework with relevant terminology standards and controlled vocabularies to create rich and highly portable metadata records. Project cataloguers created metadata descriptions using custom Web-accessible editors that interface with a MySQL database. The ETC's MySQL image database resides on a Unix (Linux) Server running Apache Web server software and is used for storing and delivering Dublin Core-compliant metadata records as well as linking them to associated image files.

### **3.2. Developing ACVA: Text Transcription, Markup, and Manipulation**

Following the completion of the digitization phase, the project team turned its attention to transcribing and marking up the text of the documents. A recent MA graduate at UNB, Dorothy Bennet, transcribed all Winslow letters that had not already been treated in W.O. Raymond's 1901 Winslow Papers, A.D. 1776-1826. Sandra Barry, a Halifax-based researcher, transcribed all of the McQueen letters. All documents were transcribed and encoded following applicable industry standards for primary source textual documents. All transcriptions were originally keyed in transcriber-selected word processing file formats. As part of the transcription process, a number of editorial conventions were applied in order to automate scheduled base-line text encoding. For example, square brackets [ ? ] and the question mark were used to indicate unclear text. When text was entirely illegible, transcribers or proofreaders would insert [gap] to denote the illegibility. Additions were noted using one of two conventions. Where added text was inline [add: N. York "inline"] was used. [add: N. York "sup"] was used to indicate when super linear text was added. This was done consistently to allow for automated markup.

Once transcription was completed, all texts were encoded in the eXtensible Markup Language (XML) in accordance with the Text Encoding Initiative (TEI) Document Type Definitions (DTD). TEI is an internationally developed markup scheme currently expressed in XML for creating, interchanging, and representing simple and complex electronic texts.<sup>[4]</sup> For this project, baseline level encoding was automated using the PERL programming language. Scripts were written to match and map textual structures and editorial conventions in the transcriptions to TEI textual and structural elements. Baseline encoding included the TEI Header, main letter-based structural elements (such as salutations and closings), additions, deletions, gaps, lineation, paragraphs, and super-scripted text. All texts were validated before

and after project encoders received files. On completion of automated baseline-encoding, project encoders used XMLSpy software to proof and edit the results, encode person and place names, as well as any structural and textual elements that may have been missed.<sup>[5]</sup>

Once encoding was complete the texts had to be delivered to the web. This was done using the XSL transformation language, an associated XML data standard. The default reading of the texts is the diplomatic version but, because the texts were normalized silently, they can be read in a normalized/regularized mode. Readers can also select to read the text with or without transcription and biographical notes, as well as with or without lineation. In simpler terms, readers can choose to read the transcription as it was actually written on the page, or they can select to read the normalized transcription, in which spelling mistakes have been corrected and abbreviations expanded in order to make the text more accessible to twenty-first century readers.

Similar to image storage, document storage is likewise provided by a MySQL database, which is loaded from XML files parsed with a Perl XML SAX parser. Database fields are populated with information parsed from XML files. At the document (text) level, information about the document's creation date, spatial and temporal coverage, and source description are recorded. Documents are stored as one or more XML objects (e.g., a page, a paragraph, etc.), with each XML object being associated with one parent document. An XML object stores information about its source file and the XML encoded portion of the parent document represented by the object. Each object also stores several text fields derived from its source XML (e.g., TEI header, diplomatic full text,) to facilitate searching. Moreover, each XML object references one or more images of its parent document. Each image is associated with its referencing XML object, its storage format(s), and the name and media type of its storage location.

Documents may be located by browsing indices organized by document title, date, author and recipient. To locate a specific document, a search index is generated from XML files that locates and scores individual words in each document, in one or more search classes (e.g., TEI header, full text, person names, place names). The search index is then used to retrieve documents matching search text in the selected search class.<sup>[6]</sup>

### **3.3. Developing ACVA: Learning Resources**

Once the imaging, transcription and markup of the collections was completed, the project team developed a means of teaching Canadian history to schoolchildren, using the electronic Winslow and McQueen collections as windows into the past. The end results of this work are the ACVA Learning Resources sites, created by instructional designer Martin L'Heureux. Both the Winslow Papers (<http://atlanticportal.hil.unb.ca/acva/en/winslow/learning/home/>) and the McQueen Papers (<http://atlanticportal.hil.unb.ca/acva/en/mcqueen/learning/home/>) have Learning Resources sections. Two main goals underscored the Learning Resources' development. First was to provide students in grades 9-12 with an easy introduction to the large amount of material held in the parent Atlantic Canada Virtual Archives. Second was to assist teachers by illustrating ways to use the Atlantic Canada Virtual Archives in classroom instruction.

All Web pages in the ACVA Learning Resources Web site are valid XHTML, with all formatting performed

through cascading style sheets. The ACVA Learning Resources Web site has been designed so that it is fully useable by older non-CSS Web browsers. Although client-side scripting is accomplished with JavaScript, the ACVA Learning Resources Web site has been designed to lose none of its essential functionality for non-JavaScript Web browsers. Server-side scripting is accomplished with PHP.

All text in the ACVA Learning Resources was created in ASCII format, then marked up in XHTML for Web delivery, with the following two exceptions:

1. The narrative content in the "Snoop" section was marked up in XML then delivered using two different methods: XSLT and Macromedia Flash.
2. The lesson plan, student handout and assessment criteria in the "Teachers" section adapted to fit the standards set forth by the National Library of Canada.

All interactive pieces of the Learning Resources were created in Macromedia Flash 6 format. Moreover, the project team was careful not to design the Learning Resources page in a vacuum. Lead instructional designer Martin L'Hereaux conducted a number of workshops with Fredericton-area high school students. At these meetings, students offered input on the design and scope of the Learning Resources, and this section of the ACVA website has been much improved with their help. One aspect of the Learning Resources page that all students enjoyed was the quill pen simulator, a Flash application that taught them the trials and tribulations of writing with eighteenth-century quill pens. This interactive learning tool is only available on the Winslow side of the ACVA site.

The development of the Learning Resources proved to be a significant challenge. To design something, especially something educational, that would be useful for a thirteen year old high school freshman while still being able to hold the attention of an eighteen year old high school senior was a difficult task. As such, this section of the site may be a little too advanced for some junior students, and it may not be "flashy" enough to hold the attention of older ones. Still, this should not minimize the effectiveness of the Learning Resources as an effective pedagogical tool. In fact, even senior academics have become enthralled with the quill pen simulator!

## **Conclusion**

Currently, the Atlantic Canada Virtual Archives holds only the McQueen Family Papers and the first few volumes of the Winslow Family Papers. It is expected that the number of collections held on the site will grow over time. For next year, an incoming MA student is planning to digitize, transcribe, and annotate an archival collection from Prince Edward Island as part of his degree requirements. This will not only add one more collection to the site, but it will also give the site greater regional cohesion by having one collection from each of the Maritime provinces. In a years' time, it is expected that only Newfoundland will not be represented on the website. Despite this lack of representation from "the rock," the Atlantic Canada Virtual Archives provides a vibrant learning environment for the general public, for scholars, and for students from grades 9-12.

## **Notes**

- \*   The author would like to acknowledge the assistance of the University of New Brunswick's Electronic

Text Centre in the writing of this paper. Special thanks to Lisa Charlong, Marg Conrad, Jason Nugent and Jennifer Whitney.

- 1. It is anticipated that future collections will be digitized and added to the ACVA site as funding becomes available.
- 2. Letter from Ward Chipman to Edward Winslow, 29 July 1783, Winslow Family Papers, Vol. 2, No. 113, University of New Brunswick Archives.
- 3. Only eldest daughter Jane did not enter teaching, perhaps because she was incapacitated following a fall from a wagon during her childhood and suffered ongoing bouts of mental illness throughout her life.
- 4. Information about TEI and TEI-based projects is available at: <http://www.tei-c.org/>.
- 5. An online copy of the project's encoding manual is available at: <http://atlanticportal.hil.unb.ca/acva/en/docs/manual/encoding.htm>.
- 6. A more detailed examination of the ACVA text transcription process can be found at: <http://atlanticportal.hil.unb.ca/acva/en/winslow/about/project.php>.