

The Atlantic Canada Portal

<http://atlanticportal.hil.unb.ca/>

Lisa Charlong, Assistant Director, **Electronic Text Centre**, Harriet Irving Library, University of New Brunswick, Canada.

Margaret Conrad, **Canada Research Chair in Atlantic Canada Studies**, History Department, University of New Brunswick, Canada.

Keywords

Atlantic Canada, *Acadiensis*, McQueen, Winslow, TEI, Humanities Computing

Introduction

A bilingual, multilayered website, the Atlantic Canada Portal is designed to explore the use of communication technology to support research related to the Atlantic Provinces of Canada. Launched in the summer of 2004, the Portal is a collaborative effort of the Canada Research Chair in Atlantic Canada Studies and the Electronic Text Centre, both based at the University of New Brunswick.

The Portal serves as a gateway to information on the Atlantic region for the benefit of researchers, students, and the general public. To this end, it serves as a digital library of primary and secondary sources, including bibliographies, theses, archival documents, reports, and teaching aids. It also supports communications among scholars by hosting virtual communities of learning, posting e-prints, and maintaining an electronic mailing list. While academic researchers, broadly defined, are the main audience for the Portal, we increasingly find that our site attracts a diversity of people interested in aspects of the Atlantic region's history and culture.

Atlantic Canada Bibliography

One of the first projects undertaken by the Portal team was to develop an online database of bibliographies published in *Acadiensis: The Journal of the History of the Atlantic Region*. The online version of the bibliography, encoded in XML and in accordance with the Text Encoding Initiative (TEI) Guidelines, enables text processing, complex searches, accepts titles that our users notice are missing, and permits us to correct the mistakes that invariably creep into any large database. Time and money permitting, we will also develop a subject index that will make the bibliography even more efficient in tracking topics.

The text files for the first segment of the bibliography (post-1939) were converted from word processing formats to XML through a series of Perl scripts. The second segment (pre-1939) was

converted to electronic format using OCR software, **TextBridge** and **Adobe Capture**. With their complex structure and punctuation, article entries were the most difficult to convert to XML. Nevertheless, the project programmer achieved 80% automation.

The project used the XML TEI Lite DTD from the **Text Encoding Initiative (TEI) Guidelines to Document Creation and Interchange**, with only slight modifications. TEI is an international and interdisciplinary standard used to represent texts for online research and teaching. The project team at UNB modelled part of its encoding on the Emblem Project Utrecht's application of the TEI Guidelines (see <http://emblems.let.uu.nl/emblems/html/index.html>). The project used the **XMLSpy editor** for text encoding.

After the bibliographic information was encoded in XML, a MySQL relational database was designed to store the source XML data. The database design is based on the structure of the encoded information for books, articles, and theses. An on-line administration system was developed to allow creation or editing of individual references. The administration system interface is written in HTML, using PHP scripts to generate dynamic content and to handle database interactions.

Following the creation of the database, a web interface was developed through which bibliography contents can be viewed. The interface offers basic and advanced search options, and a browse feature that provides lists of references ordered by title or author. The interface is written in HTML and, again, PHP scripting is used to generate dynamic page content. Users enter search criteria that are passed to PHP utilities that interact with the database to search the fields specified in the user's query. Where a match occurs, the source <biblStruct> XML for a particular record is retrieved. XSL stylesheets are applied to the retrieved XML and transformed using Sablotron, an XSLT, DOM and XPath processor. The transformation generates HTML output, which is then delivered to the user's web browser.

Atlantic Canada Virtual Archives

The Atlantic Canada Portal also hosts an **Atlantic Canada Virtual Archives**, which is designed to showcase some of Atlantic Canada's rich archival sources and to explore the potential of web-based research. In this phase of the project, undertaken in 2003-04, two frequently-consulted archival collections have been digitized for online delivery: the Winslow Family Papers, housed in the University of New Brunswick Archives and Special Collections, and the McQueen Family Papers located in the Nova Scotia Archives and Records Management, and in three repositories in British Columbia – the British Columbia Archives, the Nicola Valley Museum Archives Association, and the Rossland Historical Museum & Archives.

Edward Winslow (1746-1815) was prominent Massachusetts Loyalist who settled in New Brunswick following the American Revolutionary War. A prolific and gifted letter writer, Winslow offers an intimate account of the Loyalist experience of re-establishing family and community life in colonial New Brunswick. Over 260 letters written to and from Edward Winslow between 1783 and 1785 have been transcribed and the Winslow Family Papers (1695-1866), consisting of over 3600 items, have been imaged.

The McQueen family letters chronicle in often breath-taking detail the activities and relationships of a farming family rooted in Pictou County, Nova Scotia, in the second half of the nineteenth century. All but one of the six daughters and the only surviving son taught school at some point in their lives, moving around Nova Scotia and, in the case of the two youngest daughters, to British Columbia, to teach. A highly mobile family, they kept in touch by writing letters that speak to us today as remarkable documents of Canadian social history. The letters written between 1866 and 1890 have been transcribed and coded to facilitate searches and the entire collection of over 1200 documents has been imaged.

The Learning resources developed for this project by Martin L'Heureux are especially fun for students of all ages as they practice writing with a quill pen and are obliged, like Edward Winslow, to dip the pen in ink regularly to be able to continue writing. In the case of the McQueens, students travel with Jessie McQueen on the recently-completed Canadian Pacific Railway across Canada in March 1888, reading the letters and postcards that she wrote en route from Pictou County, Nova Scotia, to Nicola Valley, British Columbia, where she had been hired to teach school.

The project team imaged, transcribed, and developed interfaces for the McQueen and Winslow collections following applicable XML and imaging industry standards. (For those interested in the technical details, see <http://atlanticportal.hil.unb.ca/acva/en/winslow/about/index.php>). We hope that these archival resources become the first of many more for the site, but the costs of large projects such as these are daunting. At its height some 29 people, including 13 students, were employed to help transcribe, image, and encode the letters. Only a grant from the Canadian Culture Online Program of Canadian Heritage, and donations in time and money from a great many people, made such a labour-intensive project possible.

Communities of Learning

As well as bibliographies and primary sources, we have a pass-protected area on the Portal called Communities of Learning that facilitates team-based research and knowledge mobilization. For example, one of the Communities of Learning on the site is devoted to a team researching Health Inequities in Atlantic Canada using lone mothers as a case study. The 13 members of the research team post papers, messages, and links on the site and work together across any distance to build a bibliography or write a paper. When the project is completed we can use the site to disseminate our bibliographies and research papers to whoever wants to read them.

E-Prints and Repositories

The E-Print section of the Portal allows us to explore the potential of electronic publishing. In this regard, we sometimes raise eye-brows among print-oriented publishers. Scholarly communication via the Internet is increasingly causing the interests of the scholar and the publisher to diverge. While commercial publishing often brings in revenue, scholars rarely receive anything for publishing their articles in refereed journals and the time between submission and publication of one's work is usually measured in years. As Stevan Harnad is fond of reminding us, a refereed scholarly website can reduce that time frame considerably while also enabling "threaded" online discussions, automated indexing, accessibility to research data,

broad dissemination, and much more (<http://www.cogsci.soton.ac.uk/~harnad/>).

Scholars are now demanding that websites be institutionally supported and that attention be paid to permanence of content, versioning, security, and adherence to technical standards. The university-based repository offers scholars a way of managing and disseminating their research in digital form and archiving it for posterity. There is also potential for federating the repositories into one virtual space, a process that would greatly facilitate online searches. Part of the package is a university commitment to manage technological change in such a way as to ensure the migration of digital content from one set of technologies to another. Clifford A. Lynch, Executive Director of the US-based Coalition for Networked Information, argues that:

“At the most basic and fundamental level, an institutional repository is a recognition that the intellectual life and scholarship of our universities will increasingly be represented, documented, and shared in digital form, and that a primary responsibility of our universities is to exercise stewardship over these riches: both to make them available and to preserve them.”

<http://www.arl.org/newsltr/226/ir.html>

The team involved in the Atlantic Canada Portal Project is especially attentive to the archival demands of the new technology and will be experimenting with **DSpace** and other repository software in the near future.

Challenges Facing Portal Development

The biggest problem facing researchers interested in engaging the Internet is funding. To find the time to actually *do* research, scholars need programmers and web masters to support them. While funding agencies recognize that scientists need labs, they are less convinced that humanities and social science scholars do and so it is difficult to find the funds to make the Atlantic Canada Portal more than a passive receptor of information. We spend a lot of time chasing “soft” money to enable us to keep our increasingly skilled Portal team of scholars, students, and technicians together.

It is, we believe, vitally important that scholars in the humanities and social sciences share our knowledge and compete for user attention in cyber space. Students increasingly expect to find their sources on the Internet and scholarly research is greatly facilitated by online resources. While the cost of embedding ourselves in the digital world is great both in terms of time and money, so, too, is the rest of the infrastructure that serves the academy. Although we expect there to be difficult adjustments to make as we continue to ride this new communications technology to its final destination, we will be increasingly invisible if we do not have an institutionally-supported, user-friendly Internet presence.
