

The Southern Alberta Information Resources (SAIR) Project

Kathy Crewdson
President
NextLibrary Inc.

Ian Dew
Consultant
NextLibrary Inc.

Wendy Osborn
Director
Southern Alberta Digital Library
University of Lethbridge

Abstract

The Southern Alberta Information Resources (SAIR) Project is a collaborative online bibliography of published resources significant to southern Alberta. This paper introduces the partners, briefly summarizes the purpose of the project, describes the progress and challenges encountered thus far, and discusses the intended project outcomes and impacts.

Introduction

The Southern Alberta Information Resources (SAIR) Project is a collaborative online bibliography of published resources significant to southern Alberta. The word collaborative is to mean jointly and multi-dimensionally supportive. Partnerships have been formed between the University of Lethbridge (host and software developer) and NextLibrary Inc. (service provider and software developer). In the future, partnerships will be formed between community groups, such as the friends of libraries, historical and genealogical societies. Partnerships need to be formed between librarians and non-librarians.

SAIR is innovative in several ways. It is successfully forming working relationships among diverse groups. It will produce a bibliography that will be freely accessible on the Internet. Finally, it has produced needed software, such as a modified batch record loader and a cross-platform collector/editor.

The goal of the online bibliography project in the initial phase is to examine and evaluate techniques and procedures through a small, representative sample. Significant information about the process has been gathered. Once the pilot database has been exposed to the public for a period of time, outside interest through usage will also be gauged. A collaborative effort from participants to furnish an abundance of information resources emerging from all parts of southern Alberta is the *sine qua non* of the project.

SAIR Project Scope

The impetus for creating an online bibliography was a grant from the Alberta Historical Resources Foundation (AHRF) Heritage Preservation Partnership Program. The purpose of the grant was to update bibliographic works compiled on southern Alberta: *A Bibliography of Information Resources and Material Relating to Southern Alberta to 1970* and *Bibliography of Materials Relating to Southern Alberta 1971 to 1990*. Both bibliographies were compiled using current technology that existed at the time. This meant visiting each institution with holdings on southern Alberta; writing by hand or typing on a catalogue card, bibliographic, authorities, and holdings information; photocopying the card; and entering the data into a word processor. Our intent was to create an online bibliography using modern technology, open source software and international standards. SAIR will be among the new wave of collaborative bibliographies.

The development of networked information services, made possible by the WWW infrastructure, has enabled very large numbers of people to discover, organize, and publish information, including bibliographies. This publishing activity is fundamentally important because it structures information locally, creating a patchy network of secondary access points. In turn, these access points enable information discovery, the formation and development of communities of interest. In sum, this activity and the enabling technical infrastructure, invites bibliographies to take on a new interactive possibility.
(Hendry *et al.* 805)

SAIR bibliography will contain:

- A listing of significant information resources: books, selected periodical articles, dissertations, theses, government documents and major research reports about southern Alberta published between 1991 and the present, plus hidden collections;
- Resources pertaining to the geographic area south of the 51st parallel (between the boundaries of British Columbia and Saskatchewan) and the 49th parallel;
- Electronic and fixed formats: databases, multimedia, video, audio, and print materials;
- Bibliographic information on the region and its communities.

SAIR bibliography is:

- Regional by design because it serves users in southern Alberta;
- A time saver since it concentrates in one source hard to find local resources for a variety of users;
- A contemporary reference research tool for corporations, businesses and institutes.

Objectives

In May 2003, the following objectives were defined:

1. To develop an inventory of significant contemporary and historical information resources which inform research and inspire new scholarship in various aspects of southern Alberta;

2. To facilitate access to significant information resources of libraries, historical and genealogical societies, and institutes of higher learning through input of their holdings and other metadata.

Process

Marketing engages interest and communicates needs from potential participants. In 2003, the authors began a marketing campaign with two printed bibliographies under each arm and visited researchers, genealogists, librarians and curators in roughly 100 institutions in southern Alberta. The intent was to harness the experience of individuals to help design the bibliography in terms of selection (what was to be included), arrangement, and design of a simple, intuitive interface for entering records into the final product. It was also felt that the more input at this personal level, the more likely the interest.

The responses from individuals varied from very keen to very confused. Therefore, detailed notes were recorded. In addition, a web-based newsletter was published to provide progress reports, and to facilitate discussion on what topics would be considered "significant". This proved to be a thorny issue. In the context of this project, it was determined that significant shall mean those topics of importance to the region of southern Alberta, including community, water, environment, border issues (related to the proximity of the international border) and health.

The wild rose was adopted as an emblem for SAIR since it is widely recognized in Alberta as a symbol representing heartiness and independence (Figure 1).

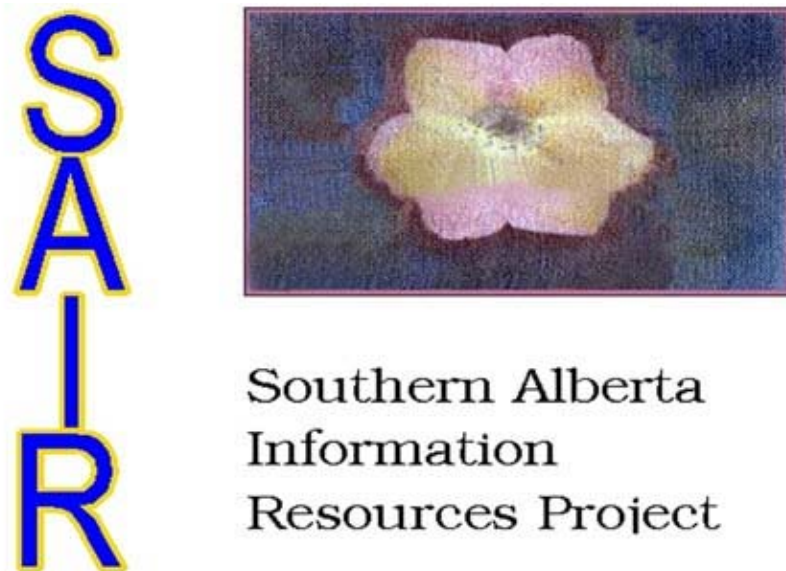


Figure 1 - SAIR Wild Rose Emblem

The approach will be to systematically load records from sources of bibliographic records for works published during the years 1991 to the present. Bibliographic records are imported in ISO 2709 (MARC) format and displayed in ISO 15836 (Dublin Core) format,

making SAIR compatible with existing and future library systems. Test searches have been made in Library and Archives Canada's AMICUS Database and a number of federated searches using ISO23950 (NISO Z39.50) compliant software. Regional bibliographic databases, including Calgary Public, University of Calgary, and University of Alberta were searched selectively. At this point, there are files of bibliographic information remaining to be loaded into the database: 2,391 bibliographic records in MARC format, published from 1991 onwards.

Since records can be added, edited and deleted using any web browser, any library or other organization, such as archives, historical and genealogical society, can contribute to SAIR in future phases. Participants have different levels of knowledge of library standards. Therefore, an experienced cataloguer will perform the final editing of records. This makes SAIR a system for true collaboration among librarians and non-librarians. Specialized search or browsing features are implemented by utilizing open source software depending on participant suggestions. For instance, the user can browse any one of nine categories: Land, Water, Environment, Community History, Community Present, Spirit, Fine Art, Economy and General.

Once a critical mass of documents on a specific topic has been loaded, users will be asked to contribute records for significant works. Critical mass is defined as the point in each user's research that SAIR becomes worth consulting. This will occur when there are more relevant records in SAIR than another source. Users can then provide records in a file in MARC format or directly using J-MARC editing software. The best-case scenario for collecting records from the region would be collaboration from every user to electronically load records to the SAIR site.

Major online resource collections to be searched include:

- Public libraries
- Alberta Legislature Library
- Archaeological Survey of Alberta Library
- The Alberta Library
- Alberta government departmental libraries
- Federal government libraries, such as agricultural research stations
- Library and Archives Canada
- Academic libraries, located at regional universities and colleges

Regional resource collections that are not available online at present include special libraries such as the Rocky Mountain Archives in Banff and the Warden's Library at Waterton Lakes National Park. Other regional resource collections to be searched include museums, historical societies and archives.

Technology

Another important intent of SAIR is to use modern technology in conjunction with accepted standards to achieve objectives. The platform is based on international metadata

standards Dublin Core and MARC to promote interoperability. In terms of software, the project uses and develops open source software out of the conviction that moving towards open systems is beneficial in opening resources to the public.

Greenstone

Greenstone Digital Library Software (Witten and Bainbridge), an open source product, is used because it is highly suitable for this project. Greenstone and any collections built with it can be customized in many ways. For example, modules that support the inclusion of new data types can be added easily to the main software base. A collection can be customized to provide different views of its contents and different searching and browsing mechanisms

The product enjoys an established funding base, has been utilized for several years by the library and archives community, and is updated frequently so that it is considered reliable. In addition, use of the software is free of the typical licensing restrictions imposed by commercial vendors, and the source code is freely available. This lowers the total cost of ownership for licensees. It is not desirable to be locked into a particular vendor solution, including their design path, as well as implementation and fixed timeline. Greenstone freely allows input so it can be tailored specifically to the community development and sharing aspects of this project.

Therefore, the intent is to maximize potential for communication and substitute a system for human labour. In order to make Greenstone more able to incorporate records, two pieces of software have been co-developed by members of the University of Lethbridge and NextLibrary. The loader software will take ubiquitous MARC records from any standard library system and load them to SAIR. The J-MARC editor will be used to create new records that can then be added to SAIR by using the loader software.

J-MARC

J-MARC (Osborn *et al.*) is a tool for creating and editing MARC records that is cross-platform, simple to use, and can be used with any digital library that supports the MARC format. J-MARC allows a community of users to truly collaborate with each other.

In designing J-MARC, existing software tools for creating bibliographic records were studied. They can be classified into the following categories:

- Integrated tools, which provide full editing capabilities for MARC records, but are built into an existing system. They are not independent tools that can be **used** with any digital library software suite such as Greenstone. Such tools include BookWhere (Webclarity Software) and IsisMarc (UNESCO), Millenium (Innovative Interfaces), TeraText (TeraText Solutions), and JavalSIS (Enea).
- Standalone tools, which also provide full MARC record editing capabilities. They are completely independent and can be used with any existing digital library suite. However, they only work on Windows-based operating systems. Since people in a collaborative environment may use different operating systems, this makes sharing

information impossible. These tools include MarcEdit (Reese), AGenT MARCIt (Auto Graphics) and Visual MARC Editor (Algorithms Consultants).

- Application Programming Interfaces (APIs), which provide a set of functions for the creation of MARC tools that suit specific needs. One such tool is MARC4J (Peters), which is an API that is written in Java. Therefore, it is platform independent after it is used to implement a software tool. However, this requires extensive programming knowledge.

Most bibliographic tools do not meet the needs of a wide audience, especially organizations that exist outside a library system. In addition, most have complex user interfaces that require detailed knowledge of how the bibliographic record is organized.

J-MARC differs from existing tools in the following ways:

1. Cross-platform support. Most standalone MARC editing tools are written exclusively for Windows operating systems. They cannot work on other systems such as MAC OS X or Unix/Linux without the setup of emulation tools. J-MARC can run on all major platforms. The only requirement is installation of the Java Runtime Environment.
2. Simpler interface. Most existing MARC editing tools have complex interfaces. J-MARC keeps editing simple by only displaying one field at a time and allowing users to hide this field when not in consideration. In addition, the corresponding field name is also displayed upfront, which removes confusion about what each field code represents.
3. J-MARC is designed with built-in flexibility so that some customization is possible. In addition, the flexibility allows for additional functionality to be added easily.

Currently, J-MARC is used with the Greenstone digital library software. It will be made available as an open-source option under the terms of the GNU Public License. It works on all versions of Windows (except 3.1), MAC OS X, and many versions of Unix/Linux.

User Interface

For J-MARC to be useful to users who are not familiar with the MARC format, it is designed with a graphical user interface (GUI) that is simple to use. Only a minimal amount of information is displayed to the user at any given time. It is also designed with flexibility with respect to the list of MARC fields that the user can work with, and the labels used to identify fields. Only a small number of the many pre-defined MARC fields are used on a regular basis, while many others are used rarely. The fields that J-MARC makes available to the user are specified in a file that is read by the GUI on startup. This file contains a tag and a label for each MARC field to be included. In future, labels will be used to set templates for editing records for material of different format types.

Field tags can be reset by the user. For example:

```
100-MAIN ENTRY PERSONAL NAME
110-MAIN ENTRY CORPORATE NAME
245-TITLE
```

Tags can be changed so that they are more meaningful to the specific group of users. For example, the fields with tags 100 and 110 can be labelled as follows:

- 100-Name of Author
- 110-Affiliation of Author
- 245-Title

Currently, the indicators cannot be customized. J-MARC will be extended in the near future so that indicators can be customized.

Example Usage of SAIR and J-MARC

In this section, the functionality of SAIR and J-MARC are demonstrated by showing how a user can contribute bibliographic information for an article on Southern Alberta. The user has recently written an article on the descendants of the early coal miners in the Lethbridge area, and feels that the existence of the article is an important contribution to SAIR.

Creating the MARC Record

The user begins by running J-MARC, which is located on their computer. The following window opens (Figure 2):

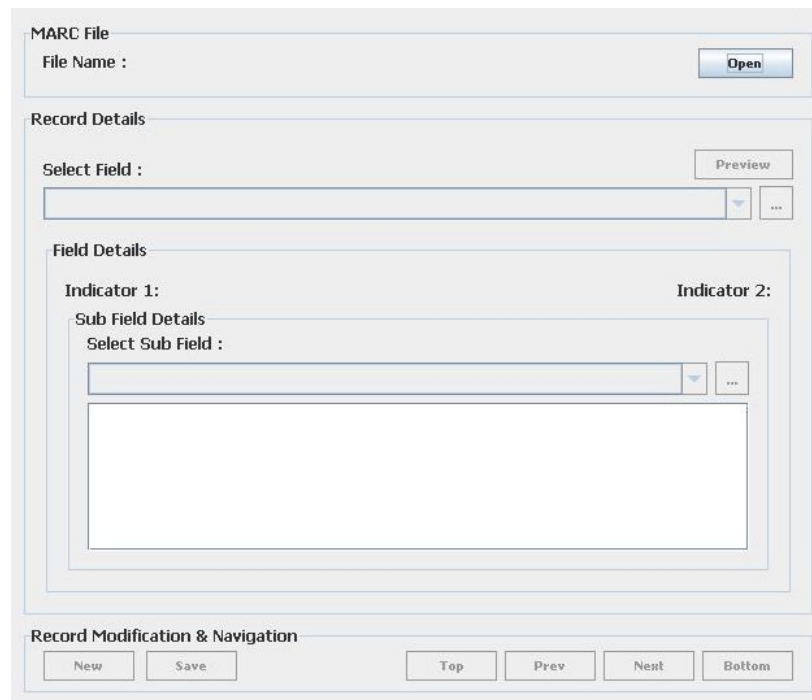


Figure 2 - J-MARC Window

The user will create a new MARC file by clicking on 'Open' in the top right-hand corner. A file dialog will appear, and in the 'File Name' field, the user will type in "famhist.marc".

After clicking on 'Open', the file dialog will disappear, and the filename "famhist.marc" will appear in the 'MARC File' area of J-MARC. Next, the user will click on 'New', which is located in the bottom left-hand corner, to add a record to "famhist.marc". Then, in the 'Record Modification and Navigation' area of J-MARC will appear the ratio 1/1, which indicates 'record 1 of 1'

Next, the user will add a MARC field. Near the top of the 'Record Details' area is the drop-down box for selecting a field. To the right of the drop-down box is a button that is clicked to produce a list of fields that are available to add to the record (Figure 3 **Error! Reference source not found.**):

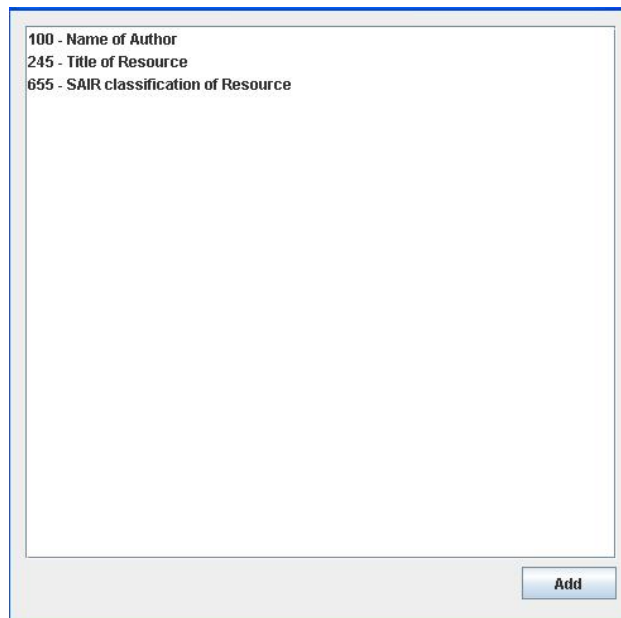


Figure 3 – List of Available MARC Fields

In this example, the fields with tags 100, 245 and 655 are available. In addition, the field with tag 655 must be added and assigned one of the SAIR categories – Land, Water, Environment, Community Present, Community History, Spirit, Fine Art, Economy, and General and Miscellaneous. Tag 655 will be used to organize records into broad subject categories for display and production of the bibliography.

The user begins by selecting '100 – Name of Author', and clicking on 'Add'. Then, the user selects this from the drop-down box. Next, a value of 1 is chosen for indicator 1 to

denote that the last name will come first. Following this, the subfield 'a – Personal Name' is selected in the same manner as '100 – Name of Author. Finally, the user enters his or her name (Figure 4):

The screenshot displays the J-MARC interface for editing a MARC record. At the top, the 'MARC File' section shows the file name 'famhist.marc' and a 'Close' button. Below this is the 'Record Details' section, which includes a 'Select Field' dropdown menu currently set to '100 - Name of Author' and a 'Preview' button. The 'Field Details' section contains two indicator dropdown menus: 'Indicator 1' is set to '1' and 'Indicator 2' is set to '#'. Underneath, the 'Sub Field Details' section has a 'Select Sub Field' dropdown menu set to 'a - Personal name (NR)'. A text input field below this dropdown contains the name 'Osborn, Wendy'. At the bottom of the window is the 'Record Modification & Navigation' section, featuring buttons for 'New', 'Save', 'Top', 'Prev', 'Next', and 'Bottom', along with a page indicator '1 / 1'.

Figure 4 – Addition of New Field to MARC Record

The user will follow the same strategy to add “Descendents of the Early Coal Miners” for ‘245 – Title of Article’ and “Community History” for ‘655 – SAIR Classification of Resource’. Finally, the user will click on ‘Save’, which is located at the bottom of the J-MARC window, to finish creating the record.

Adding Records to SAIR

After the user creates her MARC record, it will be added to the SAIR collection by visiting the Southern Alberta Digital Library (SADL) website and clicking on the image “The Librarian Interface”. After entering the SAIR project username and password, the user will click on “Launch Greenstone Librarian Interface” to display the GUI (Figure 5):

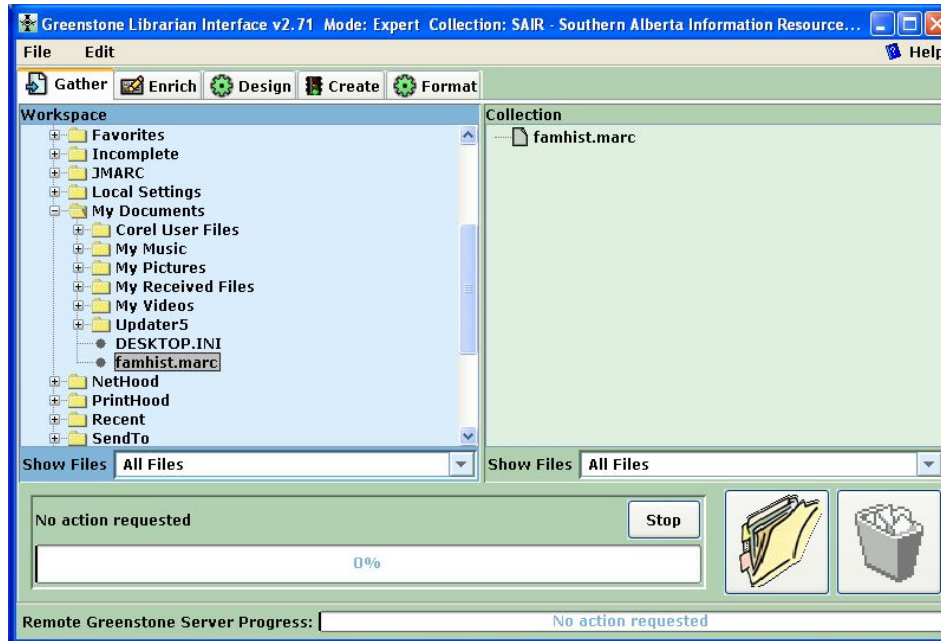


Figure 5 – Greenstone Librarian Interface

The user will begin by selecting 'Open' from the 'File' menu. A selection dialog will appear. The collection 'SAIR – Southern Alberta Information Resources' will be listed. The user will select this collection, and its configuration information will be loaded. Then, the file "famhist.marc" is dragged from the Workspace panel to the Collection panel. Next, the 'Create' tab is accessed, and 'Build Collection' is clicked to add the MARC record. Following this, 'Preview Collection' can be clicked to access the SAIR collection.

By browsing the 'Community History' link, the user can view the newly added record (Figure 6):

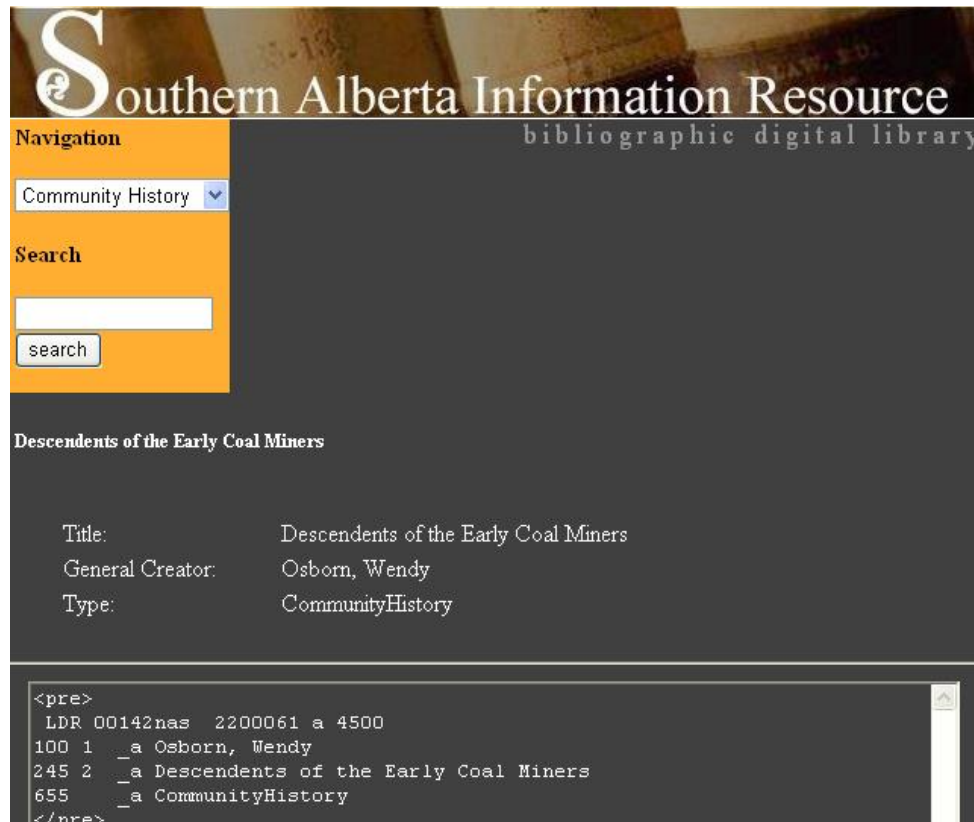


Figure 6 – SAIR Display of New Record

Two versions of the record are displayed. The first contains selected MARC fields that are displayed in Dublin Core format. The second is the original MARC record in its entirety.

Conclusions

The value of bibliography is expressed eloquently by Bill Harnum, a representative of the University of Toronto Press, "Bibliographies are really the backbone of a nation's or culture's literature. It's not too much to say that Peel's bibliography codified prairie literature for many generations to come" (Osadetz 5).

SAIR is based on a democratic concept that is intended to open the resources of the region to the general public in a way that no other initiative does. A bibliography on Southern Alberta was selected for two reasons: the authors' familiarity with the region and its literature, and the desire to revisit the bibliographic enterprise with contemporary tools.

Progress has been made in defining the scope and objectives of the project, evolving methodology and technology; however expected value and benefits are still potential. Some remaining issues and challenges are still to be resolved, specifically:

- Digital formats present a number of challenges in terms of selection and presentation. For example, how to define "published" material as opposed to archival material.
- Legal issues relating to technology such as linking and location information have emerged.
- Basic technical issues remain, such as how best to update links.
- Ownership issues related to records are more complex with consortia and involvement of quasi-commercial services, such as Online Computer Library Center (OCLC).

The level of software development is expected to be sufficient to achieve the “critical mass” of records. The next round will see tools for non library people, based on Dublin Core and Open Archives Initiative (OAI). Upon completion, SAIR will contain bibliographic, preservation, ownership, administrative and location information on documents about the region, its communities and regionally important topics. The efforts of national and provincial initiatives will be watched closely to make sure that the Bibliography will be aligned with the next generation of technology and standards.

Some important and intriguing questions have been posed, yet still remain undefined, for example, standards relating to languages and character sets. The adoption of character sets alone will present a major challenge for people involved in SAIR with major language groups represented in southern Alberta including Blackfoot and other First Nations, Japanese, Chinese, and other major Western European languages, such as Dutch, German and Italian. Greenstone is multilingual and through the use of UNICODE presents possibilities for dealing with character sets other than Latin based.

One of the benefits of conversing directly with potential participants was learning of the need for constructive action in the management of special collections, such as pamphlets, images, papers and manuscripts. Finding aids are becoming too costly to maintain and are accessible only on-site. SAIR is a perfect resource for getting these hidden collections out into the open.

SAIR is the product of four years of traveling, emails, phone calls, loading records, reloading records – failures and triumphs. The tools have been developed and we are ready to commence adding the details of the participants. In order to initiate Phase 2, more funding is needed, in addition to support from librarians and non-librarians. The bibliography would be a natural fit with corporations. Time will tell whether it is adopted widely enough to be the social networking experiment that is envisaged. SAIR will be an important step in a larger digitization initiative to preserve Alberta’s past – and more importantly, its present – for future generations to study and appreciate.

For further information on the project, contact Ian Dew at (807) 344-3831 or ian@nextlib.com

Acknowledgements

SAIR is the product of many diverse talents. The authors are indebted to the following people for their hard work in implementing portions of the SAIR and J-MARC software: Attila Aros, University of Lethbridge undergraduate student; Steve Fox, former University of Lethbridge student, now a technical specialist at the University of Lethbridge Library; Daljit Kaur, former University of Lethbridge student, now a software evaluator at COE Newnes McGehee, a manufacturer of wood-processing products. The authors are especially indebted to John Paterson, NextLibrary Inc. and Doug McCormack, CorMac Technologies Inc. for hosting and operating test sites for the past four years.

References and Bibliography

Progress with the project since its inception in 2003 is described in newsletters and other documents. <http://nextlib.com/projects.htm>. For the SAIR test database, Select "Greenstone Experimental Site".

AGent MARCIt™. 2008. Auto-Graphics, Inc. 13 Mar. 2008 <<http://www4.auto-graphics.com/solutions/agentmarcit/agentmarcit.htm>>.

Visual MARC. Algorhythms Consultants. 13 Mar. 2008
<<http://www.slimpp.com/SlimPPsite/vismarc>>.

Crewdson, K. and I. Dew. "The Southern Alberta Information Resources (SAIR) Project." CIDL News (Feb. 2006). 13 Mar. 2008 <<http://epe.lac-bac.gc.ca/100/206/301/lac-bac/cidl-ef/2007-09-28/cidl/040021-200-2006-02-e.html>>

Crewdson, K., I. Dew and W. Osborn. Southern Alberta Information Resources (SAIR) Project Update. Newsletter of the Library Association of Alberta July 2006:16. 13 Mar 2008
<<http://www.laa.ca/newsletter/2006/26.pdf>>

Dew, I. Bibliography of Material Related to Southern Alberta Published to 1970. Lethbridge: University of Lethbridge Information Resources Centre. 1975.

Enea, R. JavalSIS. 23 June 2006. UNESCO. 13 Mar 2008

<http://portal.unesco.org/ci/en/ev.php-URL_ID=5332&URL_DO=DO_TOPIC&URL_SECTION=201.html>.

Fattahi, R. and E. Afshar. "Added Value of Information and Information Systems: A Conceptual Approach." Library Review 55.2. (2006):132-47. 13 Mar. 2008

<<http://eprints.rclis.org/archive/00007728/>>.

Hendry, D.G., J. R. Jenkins and J.F. McCarthy. "Collaborative Bibliography." Information Processing and Management: an International Journal 42 (2006):805–25.

Millenium. 2008. Innovative Interfaces. 13 Mar. 2008 <<http://www.iii.com/mill/index.shtml>>.

NextLibrary, Inc., 2004. 18 Mar. 2008 <<http://www.nextlib.com/>>.

Odasetz, S. "Bringing Prairie History to Life." University of Alberta Express News. 9 May 2003. 13 Mar. 2008 <<http://www.expressnews.ualberta.ca/article.cfm?id=4368>>.

Osborn, W., D. Kaur, K. Crewdson, and I. Dew. "A Cross-platform Solution for Bibliographic Record Manipulation in Digital Libraries." Proceedings of the IASTED International Conference on Communication, Internet and Information Technology. Ed. R. Alhaji. Calgary:Acta Press, 2007. 193.198.

Peters, B. MARC4J. 2006. CollabNet, Inc. 13 Mar. 2008 <<http://marc4j.tigris.org/>>.

Reese, T. MarcEdit. 24 Feb. 2008. Oregon State University. 13 Mar. 2008

<<http://oregonstate.edu/~reese/marcedit/>>.

Snelson, P. "Communicating the Value of Academic Libraries." College and Research Libraries News 67.8 (2006). 13 Mar. 2008

<<http://www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2006/september06/valueacademiclibraries.cfm>>.

Sumner, T. and M. Dawe. "Looking at Digital Library Usability from a Reuse Perspective." Proceedings of the First ACM/IEEE Joint Conference on Digital Libraries. 2001. 416-425.

TeraText. 2008. Science Applications International Corporation. 13 Mar. 2008 <<http://www.teratext.com/>>.

Southern Alberta Digital Library. 2003. University of Lethbridge. 13 Mar. 2008, <<http://www.sadl.uleth.ca/gsd/cgi-bin/library>>.

Spinak, E. and A. Fernandez. IsisMarc. 23 June 2006. UNESCO. 13 Mar 2008 <http://portal.unesco.org/ci/en/ev.php-URL_ID=11041&URL_DO=DO_TOPIC&URL_SECTION=201.html>.

BookWhere. 26 June 2007. Webclarity Software, Inc. 13 Mar. 2008 <<http://www.webclarity.info/products/bookwhere>>.

Wick, D.E. Bibliography of Materials Relating to Southern Alberta, Primarily from 1971 to 1990 [s.n., s.l.]. 1993.

Witten, I.H. and D. Bainbridge, How to Build a Digital Library. San Francisco: Morgan Kauffman, 2003.