Digitization Project Tips Manual

Jessica A. Posgate
Digitization Project Facilitator
April 24, 2008

Contents

Introduction

This Project Tips Manual is designed to provide a practical guide for other organizations undertaking digitization projects. We hope that our experiences are valuable for preparing similar projects, from writing the grant proposal to launching the final product.

The Project Tips are arranged to match the different arcs in the project lifespan. Within each of these sections and subsections are contextual information, a series of tips or guiding questions to illustrate the events that make up different aspects of a digitization project, and technical information from the Grey Highlands Public Library Digitization Project.

For more details about the project contact the Grey Highlands Public Library at: agnesmacphail@greyhighlandspubliclibrary.com.

The Grey Highlands Public Library Digitization Project

Funding the Project
The GHPL proposed their Digitization Project to the Ministry of Culture who responded with funding in the form of a Library Strategic Development Fund (LSDF) grant that would afford a facilitator and a ten-month endeavour. Since extra staffing was necessary within this time span, GHPL also applied to the Ministry of Training, Colleges and Universities (MTCU) and was awarded a Job Creation Partnership (JCP) grant to hire a Project Assistant. Helping us keep within our limited budget, hosting services for the Agnes Macphail Website is provided by the GHPL's local server, Bruce Telecom (BMTS), who also sponsor the library's home website. The Digital Collection and Newspaper Index are hosted and managed by Our Ontario free of charge as part of Knowledge Ontario Services.

Recognizing that regional libraries needs to extend beyond brick and mortar services, the GHPL initiated a project in 2007 to digitize local resources and historical documents. Based on a collaborative partnership with Grey County museums, archives and publishers, the GHPL Digitization Project is two-pronged. One part of the project is a stand-alone website about Agnes Macphail that is accompanied by a Digital Collection of original materials collected from local repositories and private collections. The other part is the digitization of the *Flesherton Advance* and *Markdale Standard*, two newspapers that have served these communities since the 1880s and are now incredible resources for all manners of historians, genealogists, and interested readers.

The Grey Highlands Public Library (GHPL) Digitization Project acts as a resource for researchers, but also showcases Grey County's involvement in the cultural, political and social development of Ontario and Canada. By digitizing local, primary material using the Our Ontario tool kits, GHPL is not only dynamically contributing to the web of knowledge available through the Internet it is also making a proactive measure toward conserving and protecting heritage material in its archival form.

Proposing a Project

Because there is so much recent interest in uploading newspapers onto the Internet, funding bodies have trouble regarding these endeavours as unique, so it is necessary to showcase innovative techniques or un-exploited material to make any digitization proposal stand out during adjudication. The GHPL proposed the digitization of two local newspapers, but had secured full copyright permissions to reproduce the full text of the newspapers by their publishers, which is an unusual position. As well, the proposal included revitalizing

> *Proposing a project is a time consuming business and involves quite a leap of faith. You must present to the funding agencies everything from the biggest picture right down to the smallest details about a project that has not taken place yet. Read the Planning A Project section for more detail.*

the memory of Agnes Macphail, a local hero and first women elected to Canadian Parliament. No matter how unique or exciting a project may seem at the proposal stages, keep in mind that many projects may have to undergo more than one application process before successfully securing the necessary funding to go ahead.

Essentially, every stage of the project must be planned out in advance to make a strong proposal. Read the next section on planning for an even deeper break down of the project components. What follows here is an overarching look at the components of a proposal.

Project Outline
Mission
- What are you & your organization hoping to achieve?
- Who or what will benefit most?

Objectives
- Be concise about what you want to achieve, make your objectives measurable—this is invaluable as your organization can then measure progress but also gain some concrete statistics for planning future endeavours

Goals
- Define each increment of what you want to accomplish using concrete terms
- Align your goals with a healthy recognition of how far your resources will stretch— staff, primary material, equipment, skills, etc.

Budget
- The scope of the project determines the budget, projects could range from in-house scanning by library employees using free services like Our Ontario to a hired facilitator managing the processing of large quantities of outsourced microfilming or from-print scanning, website and database construction, etc.
- Equipment, staffing, infrastructure, advertising, primary material and reproduction costs, software, mileage, administrative costs, shipping, and miscellaneous costs should all be factored in to the proposed budget

Collaboration
- What other individuals or organizations are willing to participate and how will they strategically advance the project toward meeting its goals?
- Define the mutual benefits of the partnership—the gains should be reciprocal

Funding
- Target specific funding agents (for grants or other short term support) whose organizations are aligned with your project mission
- Identify and secure internal funding sources that may be necessary to make up for shortfalls from other sources

Planning a Project
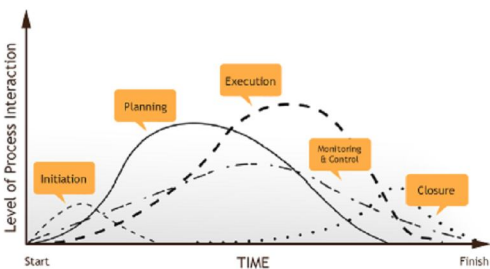
The planning stage of any project involves the most time and everything required to plan a project will have to be thought through to draft your project proposal. Planning is a worthy place to put lots of energy at the outset, but is also a constant and reflective activity throughout the lifespan of the project. The graph (right) shows the typical project lifecycle[1] and the arcs of relative time involved for the initiation, planning, execution, monitoring & control, and closure components of a project from start to finish. When planning a project of any scope, issues to seriously consider are: staff, equipment and infrastructure, desirable products, measured progress, risks and contingency plans, and collaboration.

Staffing
- What are your staffing resources? What are their skills?
- Can any internal staff be reallocated to the project? Full or part time?
- Is there a volunteer base at your organization that can be mined?
- If using external or short-term hires, budget time throughout the project for recruitment, interviews and training

Equipment & Infrastructure
- What equipment is in place, and what will you need to purchase?
- What telecommunication, scanning, computer equipment is available?
   - Your CEO will know about any existing infrastructure that can

In-house Equipment & Software

- Flat-bed scanners (e.g. HP Scanjet 5300C, Canon LiDE 30)
- Microfilm reader-scanner (Canon MS300, Capture Perfect)
- OCR software (ABBYYFineReader OCR 8.0 Professional Edition)
- Photo manipulation software (e.g. Photoshop, Microsoft Office Picture Manager, GIMP[1])
- Web-based toolkit: Our Ontario VITA and PICA toolkits
- HTML editor (MSFrontPage, Adobe GoLive, W3C HTML Validator)
- Scanner pen (CPen 20)

---

[1] Source: http://www.gnsegroup.com/application_development_process.shtml

support the requirements of processing, hosting and storing the project materials

- What are the projected consumable costs and stock necessary for the project?
  - Consider everything from envelops to printer ink to microfilm reader light bulbs
- Workspace is a crucial issue in most libraries, so consider where the project will take place and where new staff can be housed

Product Standards

Digitization work for the Agnes Macphail material was done locally and using available equipment and software, whereas the newspaper scanning was outsourced to vendors. By using the Our Ontario toolkit, the GHPL project had some established standards and guidelines to follow. In working with vendors, however, it was necessary to be able to assess and communicate product specifications to these third-party producers. Before you start, acquire basic knowledge by researching similar projects' documents and experimenting with your own equipment and materials.

- Perform a sample digitization in-house and estimate the time & expense involved for all the material you plan to digitize yourself
- Outline what you want done by your vendor and approximately when you want it done
- Get samples, financial and delivery time quotes from vendors if you plan to outsource any production—using your material

Scope Creep

The GHPL project was challenged by the newspaper production because we gave our host a huge amount of digital content in a size and format that required them to alter and adjust their workflows. As a result, we started indexing the newspaper contents from the microfilm copies into the web tool, and linked the indexes to the digital material once it was available.

During the project, we experienced all kinds of other obstacles, including long delays in hiring an assistant, family tragedies, inclement weather, and a burglary. This simply meant shuffling priorities, but never was there a gap in production.

Measuring Production

Formally measuring production with a work plan, regular reports, milestone expectations, and routine documentation are necessary steps in a well-run project. Having hard figures and facts about your expectations and how progress is made not only provides motivation during the project, they will help you learn from the project's successes and shortcomings for future endeavours. In your project plan,[2] outline a timeline in which you:

- Schedule milestone goals
- Schedule regular meetings and prepared reports that communicate production successes, failures, and issues
- Document the process as it progresses for the "post-mortem" report

Contingency Planning

There will always be mishaps, obstacles, and unforeseeable problems (i.e. That's life!). You cannot prepare for everything, but you can try to anticipate problems. Build in flexibility.

- Do a risk analysis (i.e. for every aspect of the project, what could go wrong) and contingency plans (i.e. how will you work around obstacles or rectify problems so that productivity can continue with the least amount of time/money/energy lost)
- Allow for "scope creep" (i.e. when things take longer than you'd planned) and unforeseen obstacles by overlapping production steps whenever possible so there are no gaps for any team member

---

[2] See Appendix A to see the GHPL Work Plan & Monthly Outline

Collaboration

The GHPL project was a direct result of collaboration. Working with our partners and donors we collected primary material, reproduction permissions, a wealth of personal and historical expertise, and a huge amount of support. In return, we provided the time and energy to showcase those items in digital form, promote local organizations, and provide our metadata to those organizations for their databases. Partnerships are meant to be mutually beneficial and the project manager(s) can establish this arrangement at the outset.

> *Project Partners*
>
> - Grey Roots Museum & Archives
> - South Grey Museum
> - Grey County Historical Society
> - Our Ontario
> - *Flesherton Advance* & *Markdale Standard* publishers
> - Private donors

- Communicate your intentions with either a formal document or by gaining their approval of your proposal paperwork
- Respect the partners' policies and procedural requirements, i.e. reproduction restrictions, copyright or terms of use, credit statements, etc.
- Deliver on any reciprocal favours or promises in a timely manner
- Be available and amenable to discussing new requests during the process

Place: Finding Room for a Project

Even though your digital product may result in a few Gigabytes of storage in the end, the process of creating that content will require time, space, and a set of very specific tools. Before launching new staff into the project, prepare the site, talk to your existing staff and take inventory of how your organization can manage the project. Also, be prepared—and prepare your existing staff—to accommodate new equipment, staff, and workflows within the existing environment.

- Where would project implementation ideally take place? What space is available?
- What effect will the project have on existing staff, workspace and workflows?
- What equipment (computers, hardware, software, microfilm readers, scanners, etc. etc.) is available? What will you need to acquire? Where will it go?
- How will any new equipment or workflow benefit the organization outside of the project?

People: The Project Team

Even though a "digitization project" sounds like a technological event, there are people behind every step. Knowing what your human resources are for the project will determine what can be accomplished, how, and by when.

This section is meant purely for determining how the project will be accomplished within your own organization, since you cannot control these issues for your partners, donors or volunteers.

Consider the following issues at the outset of your project planning and every time the project plan is adjusted either to meet new goals or because of staffing turnover.

Funding
Staffing costs will likely be one of the largest parts of your budget, so knowing ahead of time what the staff requirements will be is crucial.

Staff size
- How many staff do you need? Will one person be enough? Or does the project require a bigger team?
- Where is the money coming from and how much money is there?
- Can auxiliary staff be hired through any employment programs?

Required skills
- What skills are needed in the team?
- What training might be required?
- What kind of work can be done by unskilled staff or volunteers?

Project Staffing

- Project Facilitator
- Project Assistants
- GHPL CEO & Branch Heads
- Volunteers
- Our Ontario Expertise & Support

Staff Skills & Experience

- Experience with HTML, CSS, XML, W3C standards
- Scanning, photo and text editing software
- Web design and software
- Writing & editing
- Research
- Knowledge of database architecture
- Project management
- Education or experience creating digital content and metadata records
- Interest in local history and constructing digital libraries

Recruitment & Hiring

Recruitment and hiring can be done ahead of project start time, or else it will take place during the project. If hiring during the lifecycle of the project, budget time accordingly.

Consider whether you will be hiring external or internal staff. Make sure that your expectations are well laid out before taking on new staff or putting staff into new positions. This will ensure that the project team members will have a secure set of parameters to guide them in their work. Once a team is established, more dynamic or flexible approaches can be introduced, but at the outset, for the purposes of hiring, be clear about what you're looking for and how to get it.

> *Advertising & Job Postings*
>
> - Professional Associations (e.g. Ontario Library Association, Museums & Archives websites)
>
> - University Websites (e.g. University of Toronto Faculty of Information Studies, University of Western Ontario Faculty of Information and Media Studies)
>
> - Local Newspapers (e.g. *Owen Sound Sun Times*)

- Who, if anyone, can be sacrificed from inside the organization for the project?
- What skills, experience, abilities and qualifications are you looking for in your external hires? Draw up a concise job description.
- How will you advertise and recruit? Where will you find the kind of people you're looking for? What kind of budget do you have for advertising the position(s)?
- Digitization is a competitive field: How much are the people you require "worth"? How much can you afford? How else can you attract candidates to your project?
- Consider the impact on administrative workflows of new staff, external hires, or new payroll designations for existing staff.

Communication

Communication is how your project team shares their ideas, make plans and make those plans reality. Establishing a firm footing in open, productive communication is essential for any team. This is not something you can write protocols or rule books about, but if the project leader has a good sense of how communication will occur as the project progresses, they will certainly have more effective interactions over time.

- Are your team all working onsite? Or offsite? If offsite, how often will they interact face to face?
- What will be the best (most accessible) form for communication between team members? Telephone? Electronic? Print? Face to face?
- Which medium works best for what kind of communication?
- How often will there be meetings? Reports? Decide who will be responsible for the reports, for attending meetings, etc.

Process: Creating Digital Content

Locating Primary Materials

The GHPL project involved two different sets of primary collections. The first, the newspapers, are held by the library system on microfilm and many issues are also stored in print form. With permission from the publishers, we sent microfilm (positives and negatives) to two different vendors for digital reproduction.

Digitizing Newspapers

Owning newspaper microfilm does not automatically mean owning the copyright! Be sure to request permission to use the newspapers for your digitization project; otherwise, you may have to negotiate to create indexes only or some other partial reproduction.

Microfilm Quality
Positive microfilm that has been used in-house will likely be quite damaged by handling and time. If possible, acquire the negative reels or else have the negatives reproduced and use the clean positive reels for digitizing.

The second set was gathered from a variety of public and private archives to build a Digital Collection of material about and by Agnes Macphail. The variable nature of this process meant that the collecting, handling and reproducing sets of items was done on demand; however, having a general set of standards for each Digital Collection was the key to creating as much consistency as possible in the final product.

Many of the following issues about collecting, handling, and reproducing primary material will have been addressed and planned for during the proposal stages of the project. These issues will also be predetermined to some degree by your organization's holdings or those of your partners, but you may also end up exploring lesser-known repositories or putting out a call for material once you know specifically what you are looking for. Unless you have a finite collection, it will only be once you start the project that you can be very specific either about how much/many items you'll be handling or what criteria the items you handle must fit.

Collecting Primary Materials
Collection management
- What are your collection parameters? Is your collection determined by creator, theme, locale, or media?
- Where is the material that makes up the collection?
- If the material is not local, how will you gather it? Will a team member travel to other repositories?
- What condition is your material in? The quality of the original determines the quality of the scan

Private donors
- How will you reach private donors and communicate with them?
- What resources do they offer? Primary material? Memories? Historical knowledge? Etc.
- In what capacity would they like to participate? (i.e. do they want their names published or withheld, do they want copyright requests directed to them or through you, etc.)

Handling & Reproducing Primary Materials
Different media
- What types of material are you handling?
- How will you digitize it? Workflow?
- What standards have you established for each media type?

Respecting partners' handling requirements
- Special handling policies (e.g. Special collections/archives are often reluctant to reproduce items using high heat scanners and photocopiers)
- Reproduction policies (i.e. Some organizations, donors, or partners might have requirements or policies that will effect how you publish the material)

Building a Digital Collection

The Agnes Macphail Digital Collection is made up of photographs, postcards, scrapbook pages, handwritten and typed letters, government documents, newspaper and magazine articles, certificates, and illustrations by and about Agnes Macphail.

The material is gathered from two local museums and a variety of private collections. One museum specifically restricts on digital reproductions to reduce illegal replication of their archival material. We gained special permission to publish the item in a different size.

Private Donors

Individual donors were contacted ahead of time and came forward during the course of the project to provide items from their private collections. They had few restrictions except to respect the privacy of Agnes Macphail's memory by not publishing sensitive material, such as her Will & Testament.

Permissions: Reproduction & Rights Management
The GHPL project had two sets of material that needed different permissions. The newspaper digitization was made possible by the carte blanche permission given to us by the publishers to digitally reproduce the full text of the *Flesherton Advance* and *Markdale Standard*. This is an unusual license and was most likely due to the fact that

The following section about copyright is by no means comprehensive, but reflects the approach we took for the GHPL project. For the best rules, read the Canada Copyright Act!

Permission & Copyright
- Items that are published and have entered the public domain require no permission
- Unpublished writing of authors who have passed away require finding the rightful heir of the author and gaining permission from that person
- Some government documents are automatically in the public domain
- Other government documents are under copyright because their copyright date falls inside the standard 50+1 year rule of thumb
- Material that has multiple creators (e.g. a writer and a photographer) needs permission granted from each creator

both papers are privately, locally owned. Such permission may not be available for all newspapers.

Acquiring permissions for the Agnes Macphail material was more challenging because we collected many different types of items that had different creators and copyright holders.

It is a good idea to have a set of basic guidelines for new staff, but with time the rules and patterns of copyright law will become familiar. To counter the complexities of our own situation, we came up with a checklist (see Appendix C) and set of guidelines based largely on information found in Jean Dryden's book *Demystifying Copyright*.[3] In general, though:

- Determine the copyright status of every item you plan to digitize: The archive or museum or library that holds the primary material does not necessarily hold copyright!
- Ask permission to use every item you plan to reproduce, this is a good way to ensure there is no binding copyright for the item
  - o Most people are willing to extend copyright for non-profit purposes
  - o Some material is sensitive—respect the copyright holder's wishes about publishing certain material!
- Include a copyright statement about every item in your Digital Collection when you publish it
- **When in doubt, leave it out**—don't publish any item until you acquire permission

---

[3] Dryden, J. (2001). Demystifying copyright: A researcher's guide to copyright in Canadian libraries and archives. Ottawa, ON: Canadian Library Association.

Digitizing Primary Materials

Depending on what your project materials are and where you are getting them from, there will be different methods for digitizing your collection. The GHPL project included outsourcing the transfer of masses of microfilm newspaper images to digital form and the more piecemeal, ad hoc digitization of the Agnes Macphail collection materials.

Outsourcing to a vendor requires some stringent research beforehand and clear communication in order to gain the best results. Many vendors are equipped to transfer documents from microfilm to digital form, but are not all are versed in how best to handle the complexities of heritage newspaper images. Before starting negotiations with a vendor or entering into vendor selection, create a set of expectations and learn enough that you can communicate those expectations clearly. Many of us are not computer technicians, so feel free to ask the vendor to explain their methods and technology in lay terms.

- What file format do you need? Do you need any auxiliary file formats?
- How will your material be delivered? How will this fit into the workflow of loading it onto the site? Can you do it manually? Or need to hand it to someone else?
- How will your digital material be stored? Do you want to store your material in any form other than on the web? What are the advantages/disadvantages of each option?
- Issues with hosting or storage by US-based companies: Does the Patriot Act[4] affect you? If so, work with vendors who only use Canadian-based servers and databases.

---

[4] The *U*niting and *S*trengthening *A*merica by *P*roviding *A*ppropriate *T*ools *R*equired to *I*ntercept and *O*bstruct *T*errorists, or USA Patriot Act, was passed 45 days after September 11, 2001. The area that concerns librarians is Section 215, which "modifies the rules on records searches so that third-party holders of your financial, library, travel, video rental, phone, medical, church, synagogue, and mosque records can be searched without your knowledge or consent, providing the government says it's trying to protect against terrorism" (Answers.com (June 21, 2004) Patriot Act. Retrieved February 20, 2008, from http://www.answers.com/topic/patriot-act?cat=biz-fin).

Standards & Specifications

Most projects will involve manual, in-house scanning of the collection and the equipment used will often be what is already available at the organization running the project. The GHPL Agnes Macphail Digital Collection process was based on improvisation and versatility. Our collection of primary material includes photographs, letters, newspapers, journals, scrapbooks, postcards, government documents, and books. The items came from so many different sources that we performed manual scanning in the library, at other institutions and at other people's houses, and some scans were created by another organization and the digital files were transferred for our use. Once collected, the manipulation of digital files was done using a variety of desktop software (for details about this and the following digitization issues, see Equipment & Infrastructure).

Establishing digitization standards at the outset is crucial for producing consistent, high quality collections. This includes scanning qualities for storage and controlling display, metadata standards, and item identification for record creation and maintenance.

Specifications
Agnes Macphail Website:
- HTML 4.01

Agnes Macphail Digital Collection:
- stored as 600dpi TIFFs
- displayed as JPEGs

Newspapers:
- JPEG2000
- PDF of image
- OCR[1] files

Scan Storage
GHPL newspaper scans:
- ½ scans delivered on an external Seagate FreeAgent 320Gb hard drive
- ½ scans stored online at Internet Archive

Agnes Macphail Digital Collection:
- TIFFs stored on 2Gb flash key, and transferred to external hard drive

Scanning quality
- What are the best quality scans your equipment can produce?
- What file format and size is best for displaying your material, i.e. do you want users to be able to zoom in, print, or reproduce your material from the Web? Do you want to restrict users ability to do these things?

Metadata
- What schema or metadata standard makes the most sense for your collection?
- What specific fields do you need to add or suppress for your collection?
- Create a metadata application profile that describes the specifics of your choices, justifies these choices, and provides enough information that new staff can apply the schema to new material without other explanation[5]

---

[5] For example, see the GHPL metadata application profile at
http://www.greyhighlandspubliclibrary.com/AgnesMacphail/Metadata%20application.htm

> *Metadata Records*
> Dublin Core-based, 36 fields
>
> *Identification Schema*
> Items are identified by their original accession number when available, or else were assigned an alphanumerical schema. E.g. items donated by J. Burrows are identified as JB001, JB002, JB003, etc.

Identification schema
- Choose or create a system to numerically or categorically identify the items as they are added to the collection
- This identification number should be assigned both to the digital item and to its record for quick cross reference

Displaying Your Digitized Material

During the proposal and planning stages, the project team needs to decide how the digital material will be presented and managed for online display. This may be affected by the nature of the material being collected and digitized, the intended audience, the desired features that you want available with your material, or the available skills in your project team.

Online databases

The GHPL project is built primarily on the low-cost hosting and web tools solution offered by Our Ontario, which manages both the Digital Collection and Newspaper Index as databases. These databases come equipped with interactive Graphical User Interfaces (GUI), or web forms, on the "back-end" for administration and data management of the tool, and a "front end" GUI for end users of the resource. The following describes how you can manage your collection with the Our Ontario web forms:

- Newspaper Index (PICA[6]):
  - Upload JPEG,[7] PDF,[8] and OCR[9] files for each newspaper page
  - Add, change, or proof article index records
  - Add, change, or proof Birth, Marriage or Death records
  - Manage the data and online display associated with your agency within the Ontario News Index
- Ontario Images (VITA[10]):
  - Upload digital files for each item in your collection
  - Add, change, and manage the descriptive, technical and administrative records for each item
  - Manage the optional features, display and data associated with your agency within the Our Ontario collection

---

[6] Press Index for Community Access
[7] An image file format named Joint Photographic Experts Group (JPEG), after the committee that created the standard (Wikipedia).
[8] Portable Document Format (PDF)
[9] Optical Character Recognition
[10] Video Image Text and Audio

Building a Website

The Agnes Macphail Website is a standalone website built in-house and provides contextual information about Agnes Macphail and the significance of the project as a whole, but it also acts as a portal to the Digital Collection and the Newspaper Index. See Appendix B for specifications for the Agnes Macphail Website.

Planning:
- What infrastructure is available at your host organization? I.e. Can your server host a standalone website? How much will that cost?
- What is the skill set of team members—web design, web content, writing for the web, etc.
- What editing software is available? And for what computer languages? E.g. HTML, CSS, XML, XHTML, XSL, etc.

Design:
- What is the extent of your content and how does it break into sections? I.e. How do pages link from one to another?
- Use pencil and paper to draw out your design first, before you start coding
- Make the design expansible for retrospective changes and additions

Usability:
- Once a prototype site is built, get some user feedback and adjust your content to meet user suggestions
- Be prepared to adjust your site again after it goes live
- Provide contact information on the site for user feedback

Debugging:
- Some editing software can cause problems with your code and lead to design issues that will not work on all browsers
- Check your web pages in Internet Explorer, Mozilla Firefox, Safari, and Opera throughout the building process and adjust code to make pages compatible before going live
- Validate your code (see Appendix B)

Going live:
- Embed hit counters, analytics, etc. to provide traffic tracking for report statistics
- Protect your web content: consider Copyleft, GNU or Creative Commons licensing[11]
- Include links to and from host & partners' websites

Increasing traffic:
- Register your website with the major search engines
- Use meta tags and keywords to make your site more find-able
- Request reciprocal links from related web sites

Administration:
- Who will manage the site and respond to any user comments?
- Do you need an allocated email address?

Planning for the future:
- Editor software—is it licensed? Open source? What is its long-term availability?
- Create template for future page creation and editing

Storing Your Digital Material
- Once you have captured your primary material in digital form, where will the digital collection be stored? How?
- How will you store the item records?
- How can this material be accessed once it has been stored?

---

[11] Copyleft is a licensing scheme that allows authors to retain some rights under copyright while allowing users to adopt or modify content so long as they perpetuate the copyleft agreement. GNU General Public License (http://www.gnu.org/copyleft/gpl.html) and Creative Commons (http://creativecommons.org/) are examples of alternative licensing.

Promoting a Project

During the lifecycle of the GHPL project, we used push and pull communication strategies, which we hope will create an enduring interest in our resources.

Push Strategies:

- news briefs in professional publications to announce project in process and upon completion
- participation in a how-to session at the OLA Super Conference and present a poster session illustrating the project process
- distribute promotional posters and postcards by mail and electronically to announce the completion of the project lifecycle

Pull Strategies:

- invitations across Grey County to librarians, historical societies, curators, and individuals to try prototype website and provide feedback
- all promotional and presentation materials are available on the Agnes Macphail Website
- web pages and promotional publications include contact information
- website content is licensed under a Creative Commons agreement to encourage information sharing and attribution rather than restricting use
- Our Ontario toolkits are designed for interactivity, by including comment forms, mystery questions, and electronic postcard features for each item in the Digital Collection
- Project Tips Manual available for other users

Promotion does not have to wait until the project is launched. Bring people's awareness to your project during the project lifecycle and start generating a buzz about it before launching it to the general public. Target your users, both the audience and the professionals who will refer patrons to the project as a resource.

There are two types of promotion—"push" and "pull"—that work differently but just as effectively. "Push" communication describes the action of "getting the word out" about your product or organization, and methods include everything from brochures to targeted mail-outs, newsletters to phone calls, speeches to exhibits, and email to websites. Push strategies are good for gaining an audience and getting them to take notice of your product. This in turn can expand your audience and build their opinion of you through simple exposure. "Pull" strategies engage the audience and help build longer lasting relationships and interest in your communication. The trick to promoting your project for long-lasting effect is to encourage interaction with your audience by engaging them with questions, providing transferable information (e.g. How-to manuals!), providing interactive application features, or broadcasting calls for feedback.

Non-profit organizations and professionals are recognizing the value of information sharing. Developments like Copyleft, GNU and Creative Commons licensing are reflections of this movement, but so are digitization projects like the GHPL project. The best promotion we could ask for is a demonstrated collaboration between institutions. Together, we are building a greater base for our patrons and our society.

Post-Project Planning

After the formal project staff have disbanded or returned to their other duties, the maintenance of your project material and its platform still falls to the organization. Consider ahead of time what maintenance will be necessary and share as much information as possible while the project is underway so there is no jarring hand-off to the staff that remain. Offer training sessions on the maintenance issues to a core group of staff who can then train others as needed. Provide maintenance documents such as: manuals, policies, best practices, contact information, etc. so that there are print manuals to augment hands-on training (see below).

Legacy documentation
Before the project lifespan is over, the facilitator should gather together or create all the documentation necessary for any new staff to be able to comprehend every element of the project, administer or manage its content, and continue to develop the project collection without having to start from scratch.

Documenting the project as it progresses is the best way for relevant material to accumulate, but it is wise to budget a stretch of time at the end of the project lifespan when documentation issues can be given the team's full attention.

As an example of legacy documentation, the GHPL project files will include collection management documents such as:
- Metadata input guide
- Scanning guidelines
- Record and index parameters
- Storage and maintenance standards and access information
- Website files and page templates

A project management archive should include:
- Contact information (donors, team members, partners, vendors or other suppliers, etc.)
- Copyright agreements
- Partnership communications
- Project proposal
- Funding agreements
- Interim reports
- Promotional material
- A "post-mortem" report that tracks successes, issues, and solutions gained from the project experience.

Recommended Reading

Digital projects & project planning:

Inventory of Canadian Digital Initiatives http://collectioncanada.ca/initiatives/index-e.html

NINCH Guide to Good Practice http://www.nyu.edu/its/humanities/ninchguide/

Our Ontario Digitization Project Resources
http://www.knowledgeontario.ca/OurOntario/our_digitization.html

Lessons Learned: National Digital Library Competition
http://lcweb2.loc.gov/ammem/award/lessons/lessons.html

Copyright (Canada):

Canadian Copyright Act http://www.cb-cda.gc.ca/info/act-e.html

Dryden, J. (2001). *Demystifying copyright: A researcher's guide to copyright in Canadian libraries and archives*. Ottawa, On: Canadian Library Association.

Our Ontario Copyright FAQs
http://www.knowledgeontario.ca/OurOntario/our_copyright.html

USA Patriot Act:

EPIC. (2004). *USA Patriot Act H.R. 3142*. Retrieved February 20, 2008, from http://epic.org/privacy/terrorism/hr3162.html

Appendix A

Grey Highlands Digitization Project   Jess Posgate, Facilitator
Project Plan Outline   12 June, 2007

Project Description
The GHPL digitization project is designed to reproduce artefacts and documents that are unique to Grey County and to make those digitized reproductions available on the World Wide Web for researchers of all kinds.

Project Goals

➢ To digitize, upload, and publish issues of the *Flesherton Advance* and the *Markdale Standard* via the Our Ontario host site, starting with the earliest available copies and including every issue after that can be accomplished in the 8 months designated for the project

➢ To create and populate a website using local artefacts and information about Agnes Campbell Macphail that reflects her as a major Canadian historical figure whose unique character was born from her upbringing and life in Grey County

➢ To produce troubleshooting tips and a reference document for other libraries doing similar projects

Project Timeline

The attached document is a projected monthly breakdown for the Digitization Project, from June 2007 up to and including January 2008 and allowing for some overflow of administrative tasks into February 2008 (Appendix A). Without going into every detail about the workflow of each task, this outline is meant to act as a guide for monitoring the progress of each aspect of the project as it is undertaken. The document will also provide a historical document for reflection at the end of the project and a basis for more astute reporting and troubleshooting for other libraries or organizations undertaking similar projects.

Deliverables

As seen in the Monthly Outline, deliverables for this project vary in scope from incremental decisions to launching the Agnes Macphail website. Each of these and their time of delivery are dependent on many external factors (e.g. vendor negotiations) and availability of resources (e.g. project assistant). The outline, then, is considered a reasonable guideline that illustrates the variety of work and tasks involved in reaching the project goals.

The primary deliverables are as follows:

Documents:
➢ Project planning and progress reports
➢ Metadata application and indexing parameter documents
➢ MS3000 Reader-Scanner operation manual for Walter Harris Memorial Library
➢ Promotional material about the website and presence on the OurOntario.ca portal

- ➢ A Troubleshooting Tips manual to be made available upon request for other libraries and organizations

Other Resources:
- ➢ A fully functional website about Agnes Campbell Macphail
- ➢ An inventory of searchable newspaper images launched through OurOntario.ca

Potential Risks

Analysing the project for potential risks at the outset of the endeavour may alleviate potential surprises as the project progresses, and allows the facilitator and other stakeholders to consider contingency plans should problems arise (Appendix B). The GHPL Digitization Project has two major thrusts: the newspaper digitization and the Agnes Macphail website. With each of these come different sets of risks, while the project is affected by some common events that may hinder progress. Because of its built-in flexibility, there is no foreseeable risk that would require the project to be abandoned completely.

Newspaper digitization:
- ➢ Vendor negotiations are starting off slowly, indicating that tardiness may be an issue throughout the project; the number of issues uploaded to Our Ontario servers will depend upon the amount of time it takes to ship, receive, and process batches of microfilm and/or scanned print material. Any technological breakdowns we have on the ground will have a similar effect upon production
- ➢ Creating full-text searchable newspaper images will be hindered by missing paper originals or bad microfilm scans; any manual scanning or outsourcing of primary materials will also add a financial burden to the project, resulting in its deduction from another area of the budget

Agnes Macphail website:
- ➢ Hosting difficulties (i.e. ineffectual software or support) may complicate uploading some website materials
- ➢ Copyright or permissions issues may hinder availability of material for publication or reproduction

Project progress:
- ➢ Without HRDC funding to hire an assistant, the data entry, manual scanning, and promotional components of the project will fall to the facilitator or to volunteers, slowing the overall progress of the project
- ➢ While there is no reason to believe that any partners will resist, any lack of cooperation regarding access to information or other resources will skew the final product of the project and undermine the foundations upon which the project—and its funding—are built

Conclusion

On the whole, the project set out by the Grey Highlands Public Library is reasonable and modest in its proposed state. The process of bringing the project goals to fruition is bound to be organic in nature, because of the number of individuals and organizational partners involved in providing material, also because of the flexible parameters that define the inventory of uploaded newspapers and materials included in the website. Both products of the project are also bound to grow well after the project completion date, with new material being added as desired and when time is available within the regular rhythms of the library schedule. The eight months dedicated to this project in 2007-2008 will produce standalone resources that will act as solid foundations for expansion, along with documentation to instruct in-house continuation and other organizations that decide to create resources of this nature.

Monthly Outline

Month	Newspaper	Mutual tasks	Website	Extra	Deliverables
June	• Select vendor & negotiate terms and agreement • Assess manual copying and/or outsourcing possibilities based on OCR confidence levels • Inventory of paper copies	• Project plan: goals and strategies	• Survey inventory of locally held ACM materials • Define parameters for web content and purpose • Explore and determine hosting possibilities (e.g. Greenstone Digital Library) • Encoding workshop • Begin bibliography	• Staff meeting 6th • Jess away 13-27th @ Encoding Workshop • Assistant hiring?	• Project plan • Vendor selection and negotiation agreement
July	• Start shipping microfilm • Begin manual scanning and/or ship primary materials	• Define and document metadata application and indexing terms • Start indexing and metadata assignation and entry upon receipt of materials in digital form	• Conduct local interviews and photography • Collect material from local holdings • Trace and scan ACM articles and material in Flesherton and Markdale newspapers • Web site	• Assistant training? • Our Ontario meeting July 11th	• Select hosting option • Metadata application and indexing parameter documents • ACM newspaper clippings collected and scanned

			• design—get templates from web designer • Bibliography		
August	• Continue manual scanning • Shipping/receiving microfilm and digitized images	• Start uploading to digital library software	• Transcribe textual material, digitize photos, scans, etc. • Photograph and reproduce other materials • Start producing mock ups of web pages • Bibliography	• Draft scanner-reader technical manual	• Web page mock ups • Scanner-reader manual drafted
September	• Design skin for OurOntario.ca page	• Data entry—continue assigning metadata and indexing • Scanning & digitizing—continue manual and shipments • Upload materials to digital library software	• Finish collecting, reproduce and digitize local collection materials, including ACM newspaper articles • Obtain and clarify all copyrights & permissions • Start creating web pages and text • Elicit feedback*	• Staff meeting 5th • Present to Library Board* • Final user manual	• OurOntario.ca page skin • Copyrights and permissions • Scanner-reader manual complete
October		• Data entry—continue assigning metadata and indexing • Scanning & digitizing—continue manual and shipments	• Determine and design interactivity options • Clean up and finish local collection reproduction and digitization		• Final web site design • Interactivities for website • Biographical and introductory text for web site

Month					
		• Begin populating Our Ontario database (ACM newspaper stuff too)	• and uploading • Web pages & text • Finalize web site design • Solicit willing users for test next month • Bibliography		
November		• Continue all regular tasks and uploading to Our Ontario and digital library software as needed	• Prototype website for user feedback • Debug, feedback, rework, etc etc • Create administrative information for website	• Assistant & Norma: Promotional materials and advertising strategy	• User survey on web site • Administrative text for web site • Promotional material strategy and design
December	• Address storage issues and produce appropriate copies for storage • [digitization of auxiliary materials]	• Clean up metadata and index files • Adjust application and index terms documentation according to any changes	• Debug website materials (links, spelling, grammar, etc) • Permission and credit statements • Begin uploading administrative and technical text material • Bibliography formatted and uploaded	• Report materials re: grant application and project planning • Draft troubleshooting manual • Draft project report • Promotional and advertising materials—feedback, reworked and finalized • Plan for launch event of some sort	• Finalized metadata and indexing documentation • ACM bibliography • Draft troubleshooting manual • Draft project report • Promotional material finalized
January	• 188*-1940s newspapers uploaded to Our Ontario • ACM newspaper materials uploaded and linked to website • [Any auxiliary materials uploaded]		• Debug, check interoperation of website and Our Ontario • Launch website	• Print & distribute advertising and promotional materials • Library Board • Project report • Troubleshooting manual	• All newspapers (incl. ACM) uploaded to Our Ontario • Distribute promo materials • Launch ACM web site
February	• Storage of all materials: scans, print, microfilm, and digital	• Survey usability of Our Ontario and web site material, also interaction	• Assess usage • Respond to user queries and adjust accordingly	• Assistant may continue uploading metadata and/or new material	• Final Project report • Troubleshooting Tips manual complete

Appendix B

Developing the Agnes Macphail Website

Computer Language
- HTML code and displayed using CSS[12]
  - Why? We had the skills to code web pages with HTML, but we also want people on dial up internet to access the pages without horrible lag time created by complicated scripts and codes like Javascript, Flash, etc.
  - In the future, web pages written in XML[13] and using XSLT[14] for display would be the preferred approach.

Editing Software
- MS FrontPage software
  - Why? Software is available at the library
  - In the future, Adobe GoLive has been suggested as a better choice.

Validation
- Validated according to W3C[15] HTML 4.01 transitional standards
  - Why? W3C provides a free validation service[16] that ensures that your code meets World Wide Web standards, which improves the overall quality of material available on the Web

Copyright
- Protected using a Creative Commons license
  - Why? Material included on our website is not "original" per se and we hope people will use it as a basic resource without feeling as though they are infringing on our rights; asking for attribution will acknowledge the authorship and site location, increasing our visibility and validity through citation

---

[12] Cascading Style Sheet (CSS)
[13] Extensible Markup Language (XML)
[14] Extensible Stylesheet Language Transformations (XSLT)
[15] The World Wide Web Consortium (W3C) is the main international standards organization for the World Wide Web (Wikipedia).
[16] Found at http://validator.w3.org/

GREY HIGHLANDS PUBLIC LIBRARY

Appendix C

Copyright Checklist

Reproduction of anything on the web—legal or illegal—is a reality but you can reduce your project's risk of infringement by safeguarding yourself and the material you publish digitally by ensuring that you have secured copyright for all material you publish and make copyright statements perfectly clear on your displayed material.

Really basic copyright checklist:
- Item Type:
  - What category of work is the item? (e.g. literary, photograph, government document, etc.)
- Item Creator:
  - Who authored the work or is the owner of the intellectual property?
- Year of copyright:
  - When was it published/copyrighted? How long does the copyright last on the item (i.e. when does it enter the public domain)?
- Reproduction Needs:
  - How will you use the item? Excerpts require a credit line, whereas publishing or reproducing complete works requires permission from the copyright holder.
- Exceptions:
  - Are there any exceptions that permit me to use the item without gaining the permission of the copyright owner? (Some exceptions apply to single reproductions and fair use of an item, but will likely not extend to the digital reproduction and publication of material that can then be used for other purposes)

Discouraging illegal reproduction is another safeguard for your material. Some available options include:
- Digital "watermarks" of copyright information on digital images
- Generating digital material that does not reproduce well for printing or other illegal replication
- Including a citation statement so that proper citation is easy for users