

Digital Scholarship Initiative Steering Committee

Annual Report 2012-2013

Background:

The Digital Scholarship Initiatives Steering Committee is a standing committee of Library Council.

The Dalhousie Libraries provide collections, services, and facilities that expedite the access and use of all scholarly information; therefore the Committee will oversee the Libraries' planning, development, maintenance, promotion, and expansion of our digital library environment.

Membership and Meetings:

The membership of the Committee is prescribed in the Terms of Reference.

In 2012-2013 the Committee was chaired by Oriel MacLennan, representing the Killam Library. Also serving for the duration were Michael Moosberger, the University Archivist, Geoffrey Brown as Head of Technical Services, Marc Comeau as LITS Manager, Sarah Stevenson as the Libraries' Web Publisher, Allison Fulford representing the Sexton Library and Joseph Wickens representing the Kellogg Library.

After her appointment as Scholarly Communications Librarian, Heather MacFadyen joined the Committee, and served as secretary for the balance of the year.

David Michels represented the Law Library for half the year, after which time he was replaced by Mark Lewis.

The MacRae Library was represented by Elaine MacInnis and by Erin MacPherson.

Creighton Barrett, Archives Assistant, attended several meetings, by invitation.

From its inception in May 2012, the Committee met on twelve occasions, roughly monthly; agendas and minutes were appropriately filed.

Activities:

1. Terms of Reference

Terms of Reference were developed in committee, and were submitted to Library Council for consideration. After adoption, they were subsequently posted appropriately in the Share directory.

2. Inventory

After some discussion of existing Dalhousie projects at the May and June 2012 meetings, it was decided at the July 2012 meeting that a comprehensive inventory of all projects and holdings in all Dalhousie libraries, including the by-then defunct Electronic Text Centre, was indispensable to progress. This was undertaken by Sarah Stevenson, with substantial input and assistance from Creighton Barrett in the University Archives.

3. National Survey

It was also decided at the July 2012 meeting that a survey of Canadian “G-13” libraries be undertaken, in an attempt to clarify questions of staffing, policy, and platforms across the country. The “G13” list was somewhat modified, for a variety of reasons, to include Simon Fraser, York, UNB, and Carleton; the francophone Quebec institutions were excluded. Of the fifteen surveyed, we received seven viable responses. We found a great diversity of platforms in play (D-Space, Drupal, DigiTool, Fedora, etc.) and a patchwork of partial staffing as well, and concluded that there is no consensus emerging as yet.

4. “*Towards A Virtual Sixth Branch*”

In September 2012, following on from needs analysis from the Faculty of Graduate Studies concerning thesis deposit, the University Librarian requested recommendations for Dalhousie Libraries scholarship system requirements. The Committee, in consultation with relevant colleagues throughout the system, produced a report detailing inventory and recommending, in broad terms, on projected needs (hardware/storage, software, staffing and administrative) and services (self-help, intervention, and consulting). The report was presented at Library Council in January 2013 and was duly deposited in the Share directory.

5. Task Forces

Following on from the *Towards ...* report and needs analysis issues arising, it was decided in committee to establish two task forces to attempt to address matters revolving around standards and platform selection. Members of the platform task force are Geoff Brown, Marc Comeau, and Mark Lewis; their report is anticipated in September 2013.

Creighton Barrett leads the task force on standards, and is joined by Sarah David, Kevin French, Riel Gallant, and Joe Wickens in this enterprise. Some interim reports have been received to date, such as *Update on Standards Development* and *Technical Specifications for Digitization of Library and Archival Material*, as well as a schematic *University Archives Data Workflow*, which will prove most useful in moving toward common sets of standards and practices across the Libraries.

6. IR Faculty Profiles project

In the fall of 2012, the University Librarian called for development of faculty profiles in DalSpace, with a view to facilitating deposit of and access to scholarly work and research in our institutional repository. Under the leadership of Sarah Stevenson, with strong support from Heather MacFadyen as Scholarly Communications Librarian, Margaret Vail of LITS, and our three SIM interns, this project is moving forward in summer 2013, with 333 profiles constructed at time of writing.

7. Grants and Donations

The best grant award received this year went to Sexton Library to support the digitization of the Royal Architectural Institute of Canada journal 1924-1974: \$25000 was received from the Canada Council plus \$5000 from RAIC, with a further \$5000 promised for next year. Our Sexton colleagues and our friends in the Faculty of Architecture are much to be congratulated on this achievement.

8. Space

By mid-year, the Committee had secured Killam 5609 as a workroom, complete with desktop computer, scanner, and basic furniture (but no paint job), and, by the end of year, the surrounding spaces had been renovated, upgraded, redecorated, furnished and equipped to accommodate staff redeployed as per our *Towards ...* recommendations. This represents a huge leap forward on the digitization front, and we are grateful to Library Administration and to Facilities Management for their support in enabling this progress.

9. Student and Contractor involvement

We greatly benefited from SIM associates working on projects for us, in the Archives most notably; these included Zachary Howarth-Schueler and Riel Gallant.

The purpose of the former was to normalize and migrate the legacy descriptive metadata from the Archives' many MS Access databases into the Archives' new holdings management system, the Archivists' Toolkit.

The latter focussed on archives collections management and decision processes.

Their reports are included here as appendices.

We also benefitted from discussions with Dr Jamie Blustein's graduate students in the Faculty of Computer Science; these included Daniel Yule (PhD candidate), Maha al-Johani and Adel al-Hejali (MSc candidates).

Concluding Remarks:

Looking back, it is impressive how much was achieved on a variety of fronts in what amounts to a challenging new discipline in archives and librarianship. All concerned are to be thanked for their dedication, diligence, creative energy, and support. We look forward to progress in the coming year on the IR/profiles front, in resolving the platform question (for awhile, at least), on consolidating standards and practices across the system, and in developing exciting new projects.

Respectfully submitted,

Oriel MacLennan
Chair
Digital Scholarship Initiatives Steering Committee

31 July 2013

APPENDIX A

Archives Metadata Cleansing Project Wrap-up Report

January -- May, 2013

Project Summary (from project outline)

The University Archives maintains a variety of legacy metadata in various formats: MS Access databases, Excel spreadsheets, XML, etc. The purpose of this project is to normalize and migrate this legacy descriptive metadata into the Dalhousie University Archives' holdings management system, the Archivists' Toolkit.

The work will require an assessment of each set of metadata and the development of procedures to extract the metadata, transform it into EAD XML, and load it into the database for further processing. A limited-term data technician will be hired to lead the project, develop and implement a workplan, delegate data activities as needed, and ensure project goals and objectives are met. The data technician will be expected to document the methodology, procedures, and outcomes in a final project report.

Project Scope

The project dealt with the following sets of legacy data:

| Fonds or Collection | Format | Note | # Records | # Edits¹ |
|----------------------------|--------------------|-------------|-------------------------|----------------------------|
| Thomas Head Raddall fonds | EAD SGML | | 995 files 2636 items | ∅ |
| Symphony Nova Scotia fonds | MS Access database | | 2872 files | 7023 |
| Eyelevel Gallery fonds | MS Access database | | 1287 files | 5000 |

¹ Total number of unique cell-edit operations performed in [OpenRefine](#). Additional data manipulation in most cases happened in MS Access, a spreadsheet, and/or to the resultant EAD document with an XML editor.

| | | | | |
|---------------------------------------|--------------------|--|---------------------------|-------|
| imX Communications fonds | MS Access database | | 10033 files | 55760 |
| James Dinwiddie fonds | Dublin Core/HTML | | 439 files | 5940 |
| Reference Collection (MS-1-Ref) | MS Access database | Additional, significant XSLT post-processing in Oxygen XML | 4334 files 15990 items | 73387 |
| Reference Collection (MS-2-Ref) | MS Access database | | 357 files | 6153 |
| Reference Collection (MS-3-Ref) | MS Access database | | 1077 files | 24408 |
| Faculty of Medicine fonds | MS Access database | Not completed due to time constraints | ∅ (2,372) | ∅ |
| Faculty of Health Professions fonds | MS Access database | | 5654 files | 13895 |
| President's Office fonds | MS Access database | | 10021 files | 16019 |
| Board of Governors fonds | MS Access database | Not completed due to time constraints | ∅ (770) | ∅ |
| Dalhousie Cultural Activities (UA-29) | MS Access database | | 4084 files | 1072 |
| Dalhousie Art Gallery (UA-36) | MS Access database | Extracted from Cultural Activities (UA-29) database | 162 files | 17 |
| Faculty of Law fonds | MS Access database | | 3200 files | 16646 |
| Dalhousie Legal Aid Society fonds | Excel spreadsheet | Not completed | ∅ (5,332) | ∅ |

| | | | | |
|-----------------|--------------------|---------------------------|---|-------------|
| | | due to time constraints | | |
| Neptune Theatre | MS Access database | Added after project start | 10381 files | 3768 1 |
| Total | | | 54,896 files 18,626 items ² | 263,0 01 |

In total, 54,896 file-level and 18,626 item-level records were processed and added to the Archivists Toolkit.

The Neptune Theatre dataset was added and processed toward the end of the project. Three fonds' datasets comprising 8474 file descriptions were not completed within the allocated project time.

The Data Technician recorded 263,001 distinct data modifications with the primary data clean-up utility, [OpenRefine](#). This represents an average of 3.58 manipulations per file- and item-level record processed. Additional data cleanup operations took place in the contexts of MS Access, spreadsheets, and text/XML editors; however these cannot readily be quantified.

Data Validation

On numerous occasions, a summer student employee was engaged to verify descriptions in finding aids against actual files, validate arrangements and assist with records in the Archivists' Toolkit.

Data Cleanup and Migration Process

A typical dataset in this project was a Microsoft Access database, containing two tables of interest, specifically Files and Series. The structure of these tables varied slightly, but generally followed a pattern. They would be typified by the following table descriptions:

| File | SeriesTitle |
|-----------------|-------------|
| CollectionTitle | StartDate |
| BoxNumber | EndDate |
| FileNumber | Extent |

² Note: We did not tally numbers of Series- or Subseries-level descriptions added to the toolkit.

| | |
|-----------------|-----------------------|
| SeriesTitle | AdministrativeHistory |
| Sub-SeriesTitle | ScopeContent |
| FileTitle | Notes |
| StartDate | |
| EndDate | |
| Restricted | |
| Notes | |

A typical migration process involved a number of steps, each involving a particular utility; they were:

1. Microsoft Access: Create SQL query and export spreadsheet
2. OpenRefine: Import spreadsheet, cleanse and prepare data for Stead(y)
3. Stead(y): Convert data to EAD
4. Oxygen XML Editor: Clean up EAD for Archivists' Toolkit
5. Archivists' Toolkit: Import EAD and review

1. Microsoft Access

Assuming an MS Access database with (at least) the File and SeriesTitle tables described above, we would create a new query based on the join of these two tables. This may be done in query design mode, but it is generally quicker to enter SQL mode and enter the following:

```
SELECT File.*, SeriesTitle.*
FROM SeriesTitle
LEFT JOIN File ON File.SeriesTitle = SeriesTitle.Series;
```

We would run the query, save it, and export the result to a common spreadsheet format, typically MS Excel.

Note: Whichever method is used to create the query, it is advisable that a LEFT JOIN be used (INNER JOIN is the most common and default type in Access). The reason being, these databases were typically built without strongly enforced relationships between tables, and it is not uncommon to find file descriptions where the *File.SeriesTitle* field does not correspond to an actual entry in the *SeriesTitle* table. This may be because the SeriesTitle table wasn't fully populated, or due to a minor typo in the *File* table. A LEFT JOIN in this case returns all *File* descriptions whether or not they match an entry in the *SeriesTitle* table; an INNER JOIN only returns those with a match, possibly quietly discarding records that could be easily fixed in the next step.

2. OpenRefine

OpenRefine is a tool for working with messy data; it was used to do the bulk of metadata cleanup. To understand the usage of Refine, consider reviewing [these tutorial screencasts](#).

We used OpenRefine to accomplish two major of goals: (1) Fixing input errors, standardizing data formats, and extrapolating new elements; and (2) preparing data for processing with Stead.

Fixing errors, standardizing data formats, and extrapolating new elements

This was the longest and most involved step in the metadata cleansing process, and varied significantly from dataset to dataset. In Refine we examined each dataset using the tool's convenient filtering and faceting features to locate segments of data which required attention.

The most common operation performed was to identify temporal data (i.e. dates) which did not conform to RAD formatting guidelines and correct them in batches using regular expression replacement statements. (For a primer on regular expressions, try [Wikipedia](#).) Another common operation would be to facets on a field containing values from a controlled vocabulary to identify typos and misnomers, and cluster synonyms.

The full breadth of cleansing operations varied significantly depending on the schema of the dataset in question and the choices made by those who initially keyed in the data.

Preparing data for processing with Stead

This step consisted of massaging data to comply with the Stead schema (see below). To do this, we would rename some fields outright, while other new fields might need to be created based on values in existing fields. As an example, we might have two fields, "Box" and "Folder"; we would need a new field entitled "container 1 number" with a textual statement that we ultimately would like to have imported into the Archivists' Toolkit. The value for this new field, in Refine terms, would be:

```
"Box " + cells['Box'].value + ", folder " + cells['Folder'].value
```

This data processing step consists of simple operations that might just as easily be performed in a spreadsheet editor. However, as major cleansing operations were best suited to Refine, we chose to do as much of our work in this platform as possible.

3. Stead(y)

We used a modified version of Stead, an open-source Ruby-based utility for converting .csv files into Encoded Archival Description (EAD) XML. Stead is authored by Jason Ronallo and available as source code or with a web-based frontend at [steady.herokuapp.com](#).

The tool was used as a functional starting point, but numerous feature additions were required to incorporate data elements not previously incorporated by Stead. The feature-improved version is available alongside the original [on github](#).

Once installed, Stead is invoked in a shell with the `csv2ead` command. The full syntax is

```
/path/to/csv2ead \  
-c /path/to/fonds.csv \  
-t /path/to/template.xml \  
-e /path/to/dalhousie.rb \  
-p \  
-o /path/to/output.ead.xml
```

Only `-c` (.csv input) and `-o` (EAD output) arguments are required; the rest are optional but recommended.

- `-t` refers to an EAD XML template file that may be used as a starting point for the EAD that Stead will build. A template file might be an EAD file created and exported from the Archivists' Toolkit containing only fonds-level descriptive data.
- `-e` refers to a Stead::Extra file which contains post-processing instructions in a Ruby class. One is included in the extended version of Stead called `dalhousie.rb` which makes minimal changes to the document format more appropriate to the Dalhousie Archives and the Archivists' Toolkit.
- `-p` stands for "pretty" and is an instruction to include line breaks and indenting in the resulting XML file. This is highly recommended; without it, the output XML will be one long line in a file.

There are also additional options which were not used in this project.

The updated version of Stead allows a significantly extended field list, which is as follows

4. Oxygen XML Editor

The EAD XML document produced by `stead` requires some manual verification and post-processing in an XML Editor to be suitable for import into the Archivists' Toolkit.

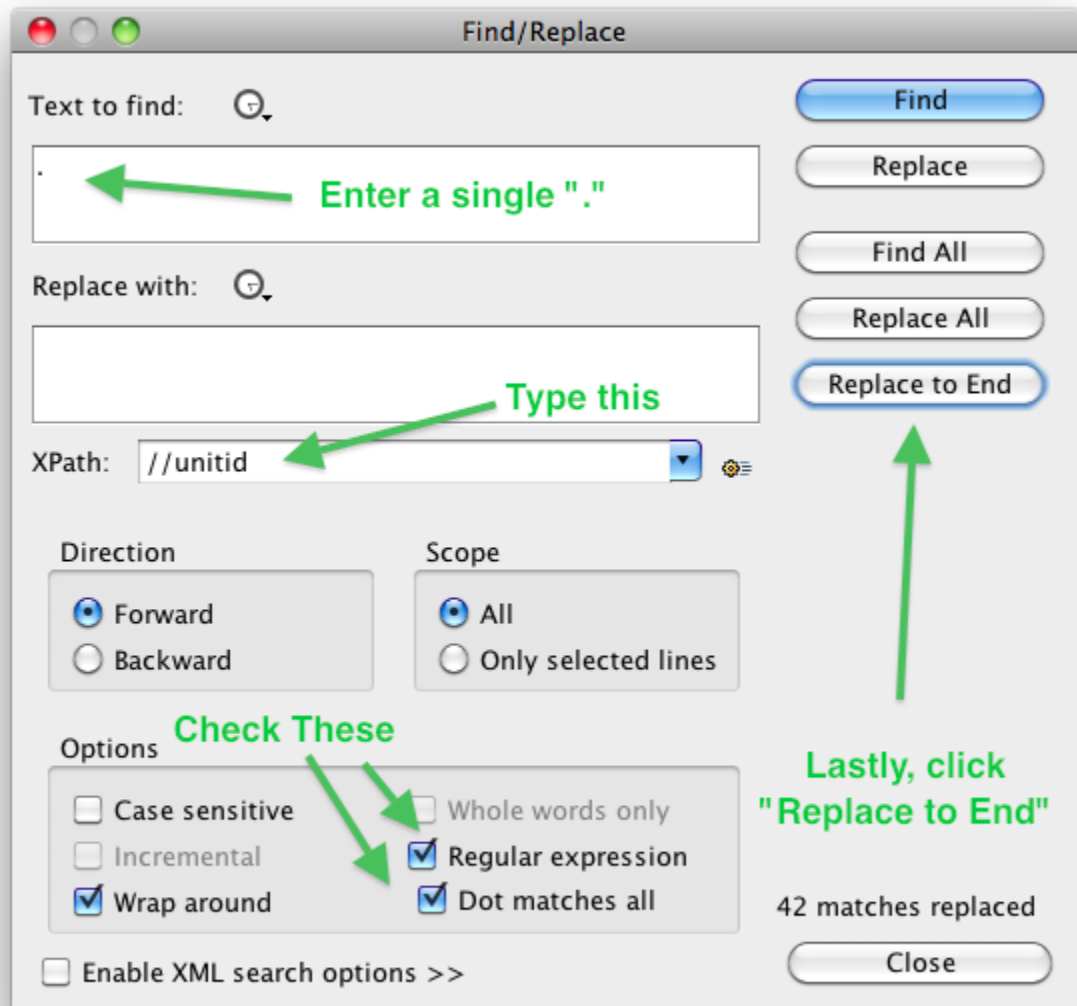
Firstly, we must verify that the document contains well-formed EAD. Some fields in EAD have formatting restrictions, but Stead will populate all fields with whatever data is contained in the corresponding column in its input file. Upon opening, Oxygen will evaluate whether a document is well formed XML that corresponds to its schema. Errors are marked below the main document window, and should be corrected.

Secondly, we must delete the `<unitid/>` tags for files, series, and sub-series elements in the document. Stead requires these ('series number' and 'subseries number' columns in .csv) in order to correctly identify finding aid hierarchy, but Dalhousie Archives does not want such identifiers imported into the Archivists' Toolkit. To remove `<unitid/>` tags (except the finding aid number):

1. Move the cursor to the <dsc/> tag:

```
98 <subject source="lcsh">Science--Experiments</subject>
99 <genreform source="rad">Textual record</genreform>
100 <subject source="local">Travel</subject>
101 </controlaccess>
102 <dsc>
103 <c01 level="series">
104 <did>
105 <unitid>1</unitid>
106 <unittitle>Correspondence</unittitle>
107 <unitdate/>
108 </did>
109 <c02 level="file">
110 <did>
111 <unitid>abcd</unitid>
112 <unittitle>A letter from James Emartz to James Dinwiddie</unittitle>
113 <unitdate type="inclusive" normal="1792/1792">August 20, 1792</unitdate>
```

- Open the Find/Replace dialogue and enter input as follows:



- Lastly, click "Replace to End". Save the document, and exit.

5. Archivists' Toolkit

This last step is basically a confirmation and touch-up step. In a testbed version of the Archivists' Toolkit, we would import the finished EAD XML document. If the import failed, we would go back to either Refine or Oxygen to correct some error in the data that prevented the Toolkit from accepting the EAD. If the import was accepted (with or without errors), we would open the newly created resource in the Toolkit for review and amendments.

Concluding thoughts

While this project specifically migrated data from legacy data stores to the particular platform of the Archivists' Toolkit, the most significant step was the cleansing of messy data, and this work may be considered platform independent. For each dataset that we processed with Refine, a complete edit history along with clean CSV and EAD documents exists, which others may use to learn from or verify the work done.

It is regrettable that we could not get to all of the datasets initially identified in the project plan. However, with the late addition of the Neptune Theatre data, we actually processed more file-level descriptions than were in the initial scope of the project.

APPENDIX B

Digital Management in Archives

Table of Contents

| | |
|--|----------|
| Part A – Profile Form | 2 |
| Part B – Decision Tree | 5 |
| Part C – Recommendations | 6 |
| Part D – Migration Plan *or* Disposition Plan | 8 |

Part A

Profile Form

This form is to be filled out to record a proper overview of the digital resource in question. Every digital resource should have a completed profile form, which can be saved as a PDF on the Digi (or S-drive?) at this location: _____

Please answer the following questions before proceeding to the Decision Tree (Part B):

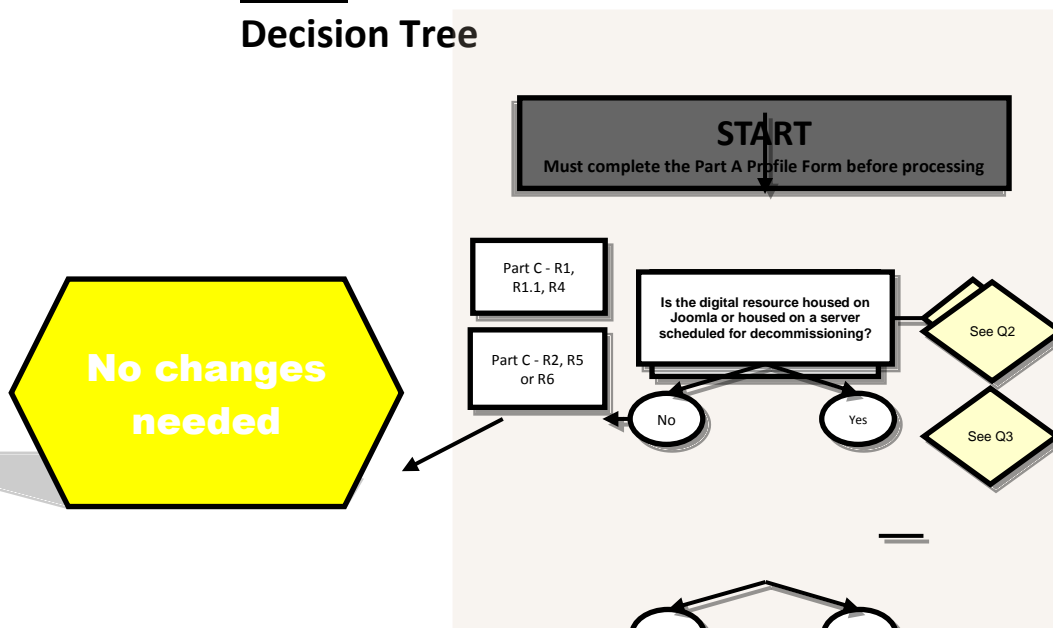
| | |
|------------|---------------------------------------|
| PF1 | Title of the digital resource: |
|------------|---------------------------------------|

| | |
|------------|---|
| | Who built the digital resource: |
| | When was digital resource built: |
| PF2 | What is the nature of the digital resource? <input type="checkbox"/> To display an entire collection <input type="checkbox"/> To display samples from a collection (exhibit) <input type="checkbox"/> To display textual information using scans of text (book(s), documents) <input type="checkbox"/> To display textual information using electronic text with no or few additional digital objects <input type="checkbox"/> Other. Please specify: |
| PF3 | On what platform is the digital resource housed? (Select more than one if necessary) <input type="checkbox"/> DalSpace <input type="checkbox"/> CQ5 <input type="checkbox"/> Joomla <input type="checkbox"/> Custom built website hosted on library.dal.ca or libraries.dal.ca <input type="checkbox"/> Off campus server <input type="checkbox"/> Other. Please specify: |
| PF4 | Does the digital resource support learning and research at and outside Dalhousie University? <input type="checkbox"/> Yes <input type="checkbox"/> No Explain how/why: |
| PF5 | Did the digital resource rely on any outside funding? <input type="checkbox"/> Yes it did, but there are no longer any obligations to the funder. <input type="checkbox"/> Yes it did, and the funders expect it to remain online in some form. <input type="checkbox"/> No it did not. <input type="checkbox"/> Unknown |
| | If yes, list any relevant conditions: |
| | Digital Objects (if applicable) Note: Digital objects are files that display information on the digital resource. These can include image files (JPEG, GIF, TIFF, PDF, PNG), text files (PDF, DOC, DOCX, TXT, RTF), audio files, video files, and more. Access copy digital objects are the files available online to the public, while the preservation digital objects tend to be larger better quality files that are backed up on an internal server. |
| PF6 | How many digital objects are there in the digital resource? <input type="checkbox"/> 1-24 <input type="checkbox"/> 25-49 <input type="checkbox"/> 50-99 <input type="checkbox"/> 100-499 <input type="checkbox"/> 500-999 <input type="checkbox"/> 1000-4999 <input type="checkbox"/> 5000+ |

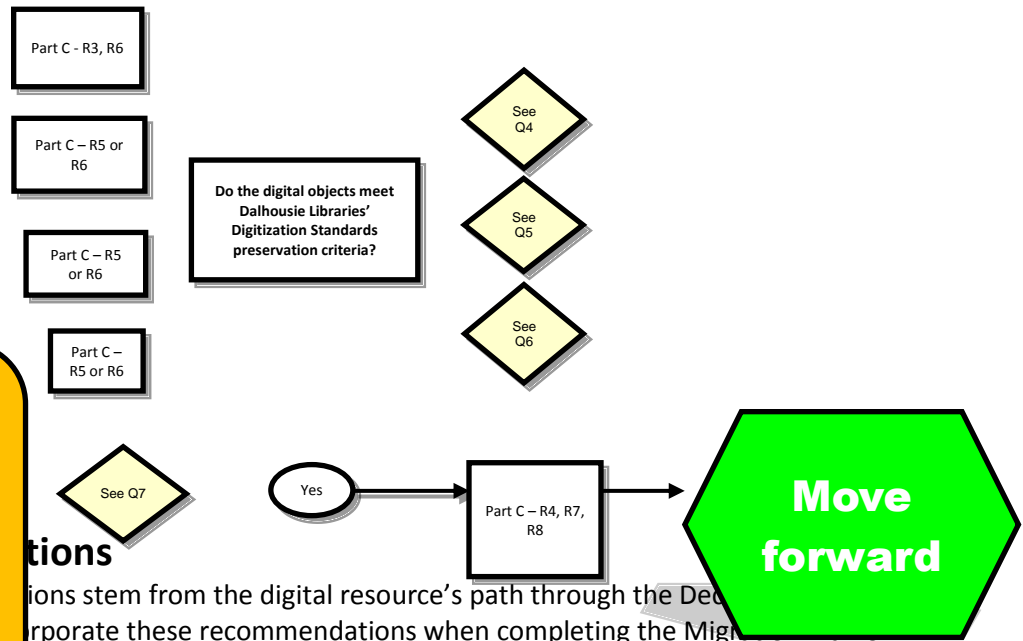
| | |
|-------------|--|
| | Specify exact number if available: |
| PF7 | <p>What file format are the digital object access copies? (access copies meaning the digital object files available on the online digital resource)</p> <p><input type="checkbox"/> JPEG (DPI:)</p> <p><input type="checkbox"/> Other image type (Specify format:)</p> <p><input type="checkbox"/> PDF (Page dimensions:)</p> <p><input type="checkbox"/> Audio (Specify format:)</p> <p><input type="checkbox"/> Video (Specify format:)</p> <p><input type="checkbox"/> Other. (Specify format:)</p> |
| PF8 | <p>What is the preservation digital object file format on the server?</p> <p><input type="checkbox"/> JPEG (DPI:)</p> <p><input type="checkbox"/> Other image type (Specify format:)</p> <p><input type="checkbox"/> PDF (Page dimensions:)</p> <p><input type="checkbox"/> Audio (Specify format:)</p> <p><input type="checkbox"/> Video (Specify format:)</p> <p><input type="checkbox"/> Other. (Specify format:)</p> |
| PF9 | <p>What is the server location of the preservation digital object files?</p> <p>Location:</p> <p>Which IT staff member is responsible for access to and maintenance of this server:</p> <p>OR</p> <p><input type="checkbox"/> These files have not been backed up on a server.</p> <p><input type="checkbox"/> These files have been deleted.</p> <p><input type="checkbox"/> These files have been lost.</p> |
| PF10 | <p>Are the digital object files on DalSpace?</p> <p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> In progress</p> <p>If yes, in what format?</p> <p><input type="checkbox"/> JPEG (DPI:)</p> <p><input type="checkbox"/> Other image type (GIF, PNG, TIFF)</p> <p><input type="checkbox"/> PDF (Page dimensions:)</p> <p><input type="checkbox"/> Audio (Specify format:)</p> <p><input type="checkbox"/> Video (Specify format:)</p> <p><input type="checkbox"/> Other. Please specify:</p> |
| PF11 | <p>What metadata fields are associated with each digital object?</p> <p><input type="checkbox"/> Title</p> |

| | |
|-------------|--|
| | <input type="checkbox"/> Description <input type="checkbox"/> Date of creation <input type="checkbox"/> Author(s) <input type="checkbox"/> Date digitized <input type="checkbox"/> Reference number/code <input type="checkbox"/> Collection/Fond/Series/Sub-series <input type="checkbox"/> Copyright <input type="checkbox"/> Edition <input type="checkbox"/> Publisher <input type="checkbox"/> Other. Please list: |
| | Technical |
| PF12 | Is there anything broken in the digital resource? <input type="checkbox"/> No <input type="checkbox"/> Yes If yes, what is the problem? <input type="checkbox"/> A broken link <input type="checkbox"/> Multiple broken links <input type="checkbox"/> Broken Flash <input type="checkbox"/> Broken Java <input type="checkbox"/> Broken XML <input type="checkbox"/> Other. Please specify: |
| PF13 | Roughly how many work hours would be needed to migrate this digital resource from its current platform to a new platform? (this includes work hours dedicated to format migration if required) <input type="checkbox"/> 1-10 hours <input type="checkbox"/> 11-25 hours <input type="checkbox"/> 26-50 hours <input type="checkbox"/> 51-100 hours <input type="checkbox"/> 100+ hours <input type="checkbox"/> Too complex to estimate. <input type="checkbox"/> Impossible/nearly impossible to migrate. <input type="checkbox"/> Highly dependent on the resources/people available. |

Part B Decision Tree



| Questions |
|---|
| Q1 – Is the digital resource housed on Joomla or housed on a server scheduled for decommissioning? Typically, this can be verified by looking at the root URL of the access point. A digital resource on library.dal.ca (as opposed to libraries.dal.ca), is most likely on a server scheduled for decommissioning (exception: DalSpace). An answer can be found in Part A - PF3. |
| Q2 – Are there any broken elements (Flash, XML, links, java, etc.) on the digital resource? Consult with the answer found in Part A - PF13. |
| Q3 – Does the digital resource support learning and research at and outside Dalhousie University? Answer can be found in Part A - PF4. |
| Q4 – Is this digital resource still relevant? Some digital resources, particularly exhibits built around current events, become irrelevant after a certain date. For example, future digital resources created for Dalhousie University's bicentennial in 2018 may be out of date once that year has passed. |
| Q5 – Do the digital objects (if applicable) meet preservation criteria of Dalhousie Libraries' Digitization Standards? These criteria are still in development and will be in place in the near future. However, good archival practices require that digital objects be scanned as TIFFs at a minimum of 600 DPI as preservation scans for the Digi server. The online access copies can be JPEGs at 200-300 DPI or quality PDFs. Some exceptions may be made for special cases (explain why in Part D: MP4, MP7, or MP10). Use the answers for PF6, PF7, and PF8 in Part A to come to a decision. |
| Q6 – Is it feasible to migrate the digital objects, along with its metadata, to a location such as DalSpace? |



Exceptions
 If this digital resource cannot be eliminated because of its importance to Dalhousie University or because of obligations to partners, see R5 in Part C to begin the process of retiring and rebuilding.

| | |
|--------------------|--|
| | <p>Recommendations needed: The digital resource is currently on a stable platform. The next step will be to complete a migration plan (See Part D) to facilitate any future migrations to new platforms.</p> |
| <p>R1.1</p> | <p>Exception: If the digital resource is on a platform that is not housed locally at Dalhousie University (e.g. ourroots.ca, etc.), ask these questions and address them in the Migration Plan (Part D):</p> <ul style="list-style-type: none"> • Does this digital resource support learning and research at and outside Dalhousie University? • Is the quality of the content of this external digital resource adequate? • Do we have the rights to the content? And if so, are there other ways to incorporate this digital resource in our discovery tools like WorldCat, DalSpace, ICA-AtoM or ALEPH? • Should the link to the digital resource remain on CQ5 digital collections webpage? |
| <p>R2</p> | <p>Broken elements (Flash, XML, links, java, etc.) can be disruptive in a variety of ways. Links can be easy to fix, but broken Flash, XML, Java, etc., may be much more difficult to repair.</p> <p>Recommendations:</p> <ul style="list-style-type: none"> • Consider how much time it would take to fix this digital resource before moving it forward. Consult an expert staff member if the digital resource is deemed worth fixing and preserving. • If it is fixable and deemed worth fixing, return to the decision tree (Part B) and choose the answer to the right for the broken elements question. Please explain in the Migration Plan (Part D) how the broken elements will be fixed (in section MP7 and MP8). • If it is fixable and deemed not worth fixing, proceed to retire or completely delete the digital resource (R5 or R6) |

| | |
|-----------|---|
| R3 | Not meeting learning or research objectives: If the digital resource is not meeting any learning or research objectives at or outside Dalhousie University, then there is no reason it should continue being housed at Dalhousie Libraries. The digital resource should proceed to either R7 or R8. Use the field in PF4 in Part A to explain why this digital resource does not meet the criteria. |
| R4 | Migrate forward: Complete Part D – Migration Form. |
| R5 | <p>Retire: Have the digital resource web archived on DalSpace as a legacy digital resource and then removed from its previous location.</p> <p>Recommendations:</p> <ul style="list-style-type: none"> • Ask Dalhousie Libraries’ Systems Group to archive the digital resource. • Put all retirement documentation in MP10 of the Disposition Plan (Part D). • If the department in charge of the digital resource plans on rebuilding the digital resource, use the Migration Plan (Part D) to explain how this would be achieved. Time, cost, expertise, equipment, re-digitizing, etc. should all be taken into consideration when building this plan. See R8. |
| R6 | <p>Complete deletion: Have the digital resource removed from the internal server (Digi) and webserver.</p> <p>Recommendations:</p> <ul style="list-style-type: none"> • Remove links from the Digital Collections CQ5 webpage. • Redirect all the digital resource’s URLs to the Dalhousie Libraries Digital Collections webpage. • Consider notifying any creators or partners. • Put all deletion documentation in MP10 of the Disposition Plan (Part D) and save it as a PDF on the Digi server at this location: Y: _____ with this file naming convention: <i>digitalresourcename_dispositionplan_date.pdf</i> • Remove online access to the digital resource. • Remove the digital resource web files and its preservation files from the internal server (Digi). |
| R7 | <p>Copy any electronic text from the digital resource and save it in a .txt file and save it in a folder entitled “Text” in the digital resource’s folder on the Digi.</p> <p>Recommendation</p> <ul style="list-style-type: none"> • If digital objects are being moved to a platform like DalSpace, consider how the text on the digital resource can repurposed. Does it need to be edited/formatted? • If digital text is a major component of the digital resource (see PF2 in Part A), consider the various ways it could be repurposed. Ask questions such as: should the webpages be copied and pasted into CQ5? Should the text be compiled into a PDF/EPUB and put into DalSpace (so it can be added to WorldCat)? Should the text just be saved on the Digi and remain dormant while it awaits being repurposed? |
| R8 | <p>Inspect the location of the preservation files: make sure the files on the Digi server are in good order. Folder structure will be addressed by the Digitization Standards sub-committee in the near future, but best practices that have been used for Dalhousie digitization projects would suggest having digital objects divided in folders by file type. Example:</p> <ul style="list-style-type: none"> • TIFF files – Folder name: TIFF • JPEG files – Folder name: JPEG • PDF files – Folder name: PDF |

| | |
|--|---|
| | <ul style="list-style-type: none"> • TXT files – Folder name: Text • The digital resource webpages could be saved (if deemed necessary) in a folder named HTML, or Web content. |
|--|---|

Part D

Migration Plan – or – Disposition Plan

A forward migration plan will make the future development of digital resources easier to manage. Using the information recorded in Parts A, B, and C, complete a forward migration plan for the digital resource in question. Each digital resource should have a completed migration plan (along with a profile form – Part A), which can be saved as a PDF on the Digi (or S-drive?) at this location:

Disposition Plan: This form can also be used as a Disposition Plan by completing MP1 and MP10, or any other relevant field(s) to the retirement or complete deletion of the digital resource.

| | |
|------------|---|
| MP1 | Title: |
| | Department responsible: |
| | Creator(s) (and partner(s) if applicable): |
| | Date of Migration Plan creation or last update: |
| | Migration Plan author(s): |
| MP2 | Explain how this digital resource supports learning and research (Found in PF4) |
| MP3 | What recommendations (R# - in Part C) did the digital resource receive after going through the Decision Tree (Part B) <ul style="list-style-type: none"> • • • • |
| MP4 | Describe what needs to be done to improve the digital resource: |
| MP5 | Describe the short term plan for this digital resource: |

| | |
|-------------|---|
| | |
| MP6 | Describe the long term plan for this digital resource: |
| MP7 | Described the digital resource's specific technical needs: |
| | Short term: |
| | Long term: |
| MP8 | Describe any potential issues users or maintainers of this digital resource might encounter (e.i. browser issues, broken elements, access issues, etc.): |
| MP9 | Indicate expected life span of this digital resource if known (e.i. 10 years, or until 2019, etc.): |
| MP10 | Disposition plan (please disregard this field if there is no plan to retire or delete this digital resource): <input type="checkbox"/> Retire <input type="checkbox"/> Complete deletion Plan: |
| MP11 | Other comments: |