Project Completion Report

Final product link:

https://digital.library.illinois.edu/collections/26bb0180-71f6-0134-1e7e-0050569601ca-5

Total hours worked:

Between 90-100 hours total. Not all hours were logged, as some email correspondence or using Skype chat were spontaneous and occurred within other work or tasks scheduled for that day. As were reviewing and fixing metadata errors on the fly.

Tasks completed:

- Item selection from an archival collection
- Participation in and observation of actual digitization process of materials
- Meeting attendance with other relevant parties and key players in the workflow
- Descriptive metadata creation for entire collection and all items contained within, metadata "profile" creation for collection (i.e. choosing metadata fields/elements most useful to potential users of this collection)
- Completion of identified reading list
- Research of other university digital collections, specifically LGBT collections
- Research outside of archives about student groups and the current LGBTRC on campus
- Participation in workflow development with Digital Content Creation Department

Overview of workflow:

The following workflow is at least the beginning of a process that future small scale digitization projects can follow between the University Archives and the Digital Content Creation department. DCC is in the process of refining this workflow and others as digital content continues to be consolidated in the new campus wide Digital Library. Projects like LGBT at UIUC will be dealt with specifically in terms of sharing information between Medusa and Archon, and then being pushed to the public facing side of the Digital Library.

1) Identify collection or collections that contain items that are good candidates for digitization. Good candidates are defined as a) in need of no or very few conservation or preservation treatments, i.e. should not be harmed by digitizing methods such as scanning, and b) projections of user audience, i.e. "this collection of items would be useful for LGBT scholars."

2) Item selection. The following should be considered at the time of item selection: project timeline, project bulk, type of material, scope of content, "message" you want the digital collection to get across (this will be an important consideration for writing the collection description later on).

3) Fill out the Digital Project Proposal online form: https://docs.google.com/forms/d/e/1FAIpQLSc95I96RJGq8byRgHvWXCxrdxeQRVtUJpxNrAYg3SWZYrVF4 A/viewform (also see Appendix A)

4) Meet with the Digital Projects Coordinator to talk about your project. Be prepared to discuss the significance of the project and the audience you hope to serve. DPC will need to know when to expect materials for scanning workflow and when items need to be ready for return to the Archives.

5) Create a CSV file for upload. This CSV will be used as a staging tool for DCC, so they have an idea of the kinds of items they can expect to be delivered. Each CSV file should contain the following columns: Record Series ID, URL for current finding aid on Archives server, Series title, Box number, Folder name or number, date (of item creation), Unique identifier for file name. The unique identifier should contain series, box, and folder number (or title) followed by the item creator, item date, and a unique name. Finally, the identifier should end with a 3 digit numeral. i.e. 001. For multi-page objects, this should be a range, i.e. 001-009. See Appendix B for an example.
6) Prepare Hollinger boxes with streamers for delivery to DCC. Streamers should be placed before the item with the 3 digit number or numbers written at the top. Arrange for transport and confirm timeline with Digital Projects Coordinator.
7) Create a metadata profile on the development side of the digital library. (For helpful hints on creating a metadata profile, see Kyle Rimkus or Trish Lampron.) Export the metadata profile from the development side as a json file and import it into the production side.
8) Once items have been digitized and ingested into Medusa, item level metadata can be created. Depending upon how many items are in the collection, metadata can be entered item by item in Medusa, or entered into a TSV spreadsheet. I completed my metadata via spreadsheet and had it uploaded. See Appendix C (created by Trish Lampron) and Appendix D. Uploading the spreadsheet once metadata is complete is the same process except choosing upload instead of download.
9) Choose item thumbnail for the collection cover and write collection level description. Once everything is set up in the development side it can be reviewed. Once satisfied, DCC can push all the content to the public side of the digital library.

This process is still under development, but these are at least the fundamental steps that needed to happen.

Problems overcome:

- There was some miscommunication about how long DCC needed to hang onto the files before they could be returned to the ARC, but I don't think that caused any real trouble.
- Illness happens. When the DPC was out for several days, things were at a standstill. Ultimately not detrimental, but frustrating when only one person can answer a specific set of questions.
- The initial TSV download for my metadata input had problems. Some individual items had been packaged together as objects as if they were multi-page items. I learned later that this was a problem with a programming script, but it was a tense few days while that was being fixed. This was ultimately part of the learning curve of a brand new work flow.
- The initial TSV upload of the completed metadata would not work for me. Kyle had to fix it and upload it for me.
- After that there was another script error that placed quotation marks around all the titles. That also got fixed on the DCC end and didn't take long at all.

Sustainability suggestions:

I'm really hoping this collection doesn't die with the end of the semester. Since we know that Siobhan's class uses the archival collection a lot, perhaps some of her students would be interested in working to get more material? Leslie Morrow at the LGBT Resource Center also expressed an interest in continuing to add to and work with the collection, and I would think Curt McKay or Pat Morey (among others) might be interested in donating material as well. Of course unpaid internships and practicum students could also work in this collection to do more scanning with DCC or create metadata. I'd really like to see

LGBT material from collections other than the 41/2/46 record series get added eventually.

Rights and privacy concerns and solutions:

Based on guidance from Anna Trammell, it was decided that the items, created by campus groups and essentially "owned" by the University, could be covered under the generic "Copyright University of Illinois Board of Trustees." If this rights statement becomes problematic in the future it can always be changed.

The RSO registration cards posed a different issue, as they contain individual names and sometimes addresses and phone numbers of those individuals. Since the cards are from the late 1970s, it is unlikely that landline phone numbers are still valid. It is also unlikely that personal addresses are valid today. After some discussion, the decision was left to the LGBT Resource Center director, Leslie Morrow. She decided the cards should be published as is, largely because RSO registration is a matter of public record. If these documents need to be redacted in the future they can be with image editing software used by DCC.

Further study:

I learned much during this short term digitization process and would like to continue studying and gaining practical experience in this niche of the LIS field. I would like to learn more about how copyright applies to different items, within and without archives, as this seems like a complex landscape in terms of digitization projects of all kinds.

I would also like to further explore hands on digitization hardware and software that is commonly used in small and large scale projects. Having been employed by the Internet Archive, I am familiar with book scanners where page turning is required in a cradle, but beyond that my experience is very limited.

Angela, the Digital Content Project Coordinator, has a job and title that I think I would really enjoy. I actually have plans to sit down with her and ask about her background, to see what professional steps I could take to pursue something similar at another university. I was a project manager for many years in the publishing field, so doing that for digital projects and workflows appeals to me.

Final thoughts:

All in all I would say that this independent study was a success, as I did what I set out to do: take a small scale digitization project from beginning to end. I started in the archives and ended online. Some steps were altered, as I initially anticipated having to build a hosted website specifically for this collection as more of a digital exhibit. But based on the current push for the new digital library, it made sense to just place the collection there. I think that will likely also make it easier to expand the material in the future. I am pleased with the end result, and I was glad that I could help Kyle, Angela, Toby, and Trish in developing their workflow by using my collection as a test case.