Centering The Margins in Digital Project Planning

Dorothy Berry

ABSTRACT

This paper uses the example of the development of a digital collections program at Harvard University's Houghton Library to provide an introduction to building programmatic and equitable collections online. Focusing on why, how, and what goes into programmatic digital collections through examples of ongoing project workflows and special projects workflows, the paper introduces methods for digital project planning that center diverse histories.

Keywords: Project planning, digitization selection, antiracism, underrepresented histories

INTRODUCTION

As archives and special collections at predominately white institutions are increasingly called to account for long-term decisions to undervalue marginalized peoples' histories, there has been an increased demand to acquire digitized material to fill in gaps. This is anecdotally evident, but also illustrated in examples like the Council on Library and Information Resources transition of the "Digitizing Hidden Collections" funding initiative to supporting only projects that highlight People of Color and other marginalized communities.[1] This paper focuses on the sorts of digital collections building that can be done looking directly at the materials on hand—whatever they may be. Regardless of the specifics of the collection, there are methods to consciously center histories that have been pushed to the margins by universities, historical societies, and all sorts of folks with the power to write the record.

Drawing on my experience building a digital collections program from the ground up at Houghton Library, I will cover three fundamental questions: why, how, and what. *Foundations*, the whys of making underrepresented histories central to digitization; *resources*, the hows of what's needed to make things happen; and *program development*, the outcomes we achieve.

Organizational structure and funding varies so widely from institution to institution that I will avoid specifics—fundamental though they are to the practicalities of getting anything done in library-land. At the same time, I will not cover grant applications and funding—fundamental though they are to getting digital projects approved at many institutions. While that may seem like I'm skipping over the most important elements at hand, this is intentional. I'm also not going to delve too deeply into post-custodial digitization and community archives work.

There are fantastic resources around digital project management, community archives, and grant applications from organizations like the Digital Library Foundation.[2] I want to acknowledge that great work has been done by colleagues who have far more expertise than I. Instead, this paper is an introductory exploration of programmatically building workflows for new digital collections that reflect the diverse stories in any collection's holdings.

FOUNDATIONS

Library professionals are widely discussing the foundational need to highlight diverse histories, especially in this current moment. While there is recurring debate in special collections about how much our work matters, the extreme limitations to physical access to collections in 2020 illustrated how core digital collections have become to syllabi and

---

[1] Kathlin Smith, "Mellon Foundation Funds CLIR's Digitizing Hidden Collections and Archives: Amplifying Unheard Voices," *CLIR News*, February 17, 2021, https://www.clir.org/2021/02/mellon-foundation-funds-amplifying-unheard-voices/.

[2] Andy Aston et al., *DLF Project Managers Toolkit*, DLF Wiki, accessed August 24, 2021, https://wiki.diglib.org/DLF_Project_Managers_Toolkit.

research plans across disciplines.[3] Where they may have previously been seen as a last resort, they are now deemed necessary.

The digitization selections that archivists and digital librarians have made in the past directly affect the selections researchers and instructors are making today. In the early months of 2020, I was part of conversations with colleagues locally and at peer institutions where we discussed advising teaching faculty to focus on materials that were already available online, and not spend too much time trying to get that perfect document digitized. These conversations called direct attention to how past decisions shape what gets taught in the classroom and published in academic journals. Not prioritizing digitization projects that center marginalized people's materials in years past almost certainly meant that those materials were not getting into classrooms in 2020.

This topical concern should be a permanent one, as we adapt more and more to online teaching and digital methodologies. As Sophia Ziegler pointed out in "Digitization Selection Criteria as Anti-Racist Action," not only does the digital corpus we create affect what materials researchers can reference, it play an outsized role on the sorts of computational research that can be done from digital collections.[4] Digital scholarship that depends on a large corpus of digitized materials is directly shaped by our choices. For instance, if we only provide access to data on early modern English books, that's the only data that will be used for research.

The audience for digital collections ranges far beyond local academics and reading room patrons. I've heard suggestions for new areas of digitization countered with, "Well our researchers are interested in ABC" or "We get most requests for XYZ in the reading room…." However, digital collections open the library door to everyone with internet access, requiring a reconception of who makes up our patron base. This broadening of access makes it even more imperative to ensure we are creating digital collections that reflect the diverse interests and backgrounds that exist in our holdings.

RESOURCES
This section will focus not on scanning specifications and budgets, but rather on thinking programmatically about whatever resources are available in a particular institution. Thinking programmatically means evaluating the resources at hand and figuring out a balance that can produce results, even if they aren't necessarily the results of your dreams. Working at Harvard University's Houghton Library has meant, for example, that funding and resources have not been my strongest concern, but the byzantine levels of hierarchy and bureaucracy that underlie any move at an institution as large and old as Harvard make staffing and time-management larger issues than I initially expected. At my institution, I manage metadata creation, project selection processes, and project management. To get a single item digitized and deposited in our institutional repository,

---

[3] Randall C. Jimerson, "Archivists and Social Responsibility: A Response to Mark Greene," *The American Archivist* 76, no. 2 (Fall/Winter 2013): 335-45, http://www.jstor.org/stable/43490358.
[4] Sophia Ziegler, "Digitization Selection Criteria as Anti-Racist Action," *Code4Lib Journal* 45 (August 9, 2019), https://journal.code4lib.org/articles/14667.

I must collaborate with multiple committees, colleagues, and departments outside my library. This can add unexpected layers of complexity to any digital project.

As I acknowledged at the outset, every institution has its own strengths and weaknesses in terms of resources, and every institution has different set-ups around staffing, funding, and infrastructure. Some recommendations in this paper will make perfect sense, and others may sound humorously out of reach. Regardless of individual circumstances, we can all do something to move towards more inclusive digital collections. Methodically evaluating the resources at hand is key. Charting out the strengths and weaknesses in terms of institutional resources will help you figure out the areas that might have room for exploration.

PROGRAM DEVELOPMENT

Using experiences I gained establishing a digital collections program at Houghton Library that centers anti-racist selection alongside other practical needs, I propose a path towards functional, programmatic collections building. As the resources section highlighted, this path is certainly not the only one and may be out of reach for many institutions. Individual elements that can be extracted will hopefully still be of use.

When I began my tenure at Houghton three years ago, patron requests generated the bulk of our digital collections requests. We receive a high volume of requests and are glad to fill them to the best of our ability as part of our public service goals. In the summer of 2020, in response to COVID-19, our workflows were shifted to bring people back on site four days a week specifically to provide free reference scans for researchers and high-quality scans for classes. Under normal circumstances, however, Harvard's Imaging Services department received, invoiced, and delivered patron requests for digitization and deposit into the institutional repository. Curators had the ability to fund and propose projects to digitize larger collections, but there was no clear understanding of the selection process or priorities. This ad-hoc digitization method meant two things: it was very difficult to get a curatorial understanding of what we had actually digitized, and that our digitization tended to reflect the research interests of our patrons. Those interests were biased both by the traditions of the academy and by a patron-level understanding of our very large collection: Harvard Library is the world's largest academic library, and Houghton Library is its largest rare books and manuscript repository.[5] Due to a history of broad collecting and interlibrary system transfers, the hundreds of thousands of materials in Houghton's unbrowsable stacks remain unknown to patrons.

The move towards thoughtful and actively selective programmatic digitization was necessary to ensure that we were meeting library-wide goals of increasing the diversity of the materials available online and fostering equity across the library. There are many opportunities to plan workflows in ways that actively pursue diverse collection representation in our digital libraries. Clarifying and codifying who decides what is

---

[5] "About Harvard Library," Harvard Library, accessed September 1, 2021, https://library.harvard.edu/about/about-harvard-library.

digitized, what criteria projects are judged by, and how we prioritize our ongoing work can lead to workflows that don't require reinventing the wheel every year. Figuring out the best approach to transition from ad-hoc to organized meant, for me, reviewing the work of colleagues at other libraries. Looking at work done by the Harvard Law Library and by the University of Minnesota library, I decided that the best option would be to create an internal mini-grant competition. Curators traditionally had the biggest internal authority on what collections were digitized, but collections knowledge is more diffuse in an organization. Catalogers, reference librarians, and programming staff all have unique areas of expertise, and the internal competition invited more people into the dialogue around digitization selection.

I created a project submission form that asked submitters to describe the collection, but more importantly to detail the benefits of digitization. The form asked about possibilities for instructional use, for potential digital scholarship projects, ways that the collection highlights underrepresented histories, and long-term preservation concerns that could be mitigated through digitization. This introduced an entirely new way of looking at digitization, and the questions definitely ruffled the feathers of some who weren't used to having to describe and defend their proposals. It also opened up completely new areas of expertise: our music catalogers in particular made proposals for materials that no one else would have thought of!

The internal application process can be divided into three steps: survey, submit, and select. The first, and most important, is to survey the landscape of your institution. Promote the submission process internally, across departments. If possible, open the process up more widely and promote outreach with faculty, students, and researchers. At the same time, create a simple submission process, asking project submitters to explain their proposal and to justify the work's value on a variety of axes. Online options, like a form, make it easy to track data.[6] The final step is to make selections. Form a rotating, interdepartmental committee of stakeholders to review and select submissions based on a shared, weighted rubric. Ideally, with streamlined organizing this can be accomplished with minimal meetings, and adding a rotation allows for diverse participation.

After multiple presentations and one-on-one conversations sharing the submission process and logic with my library, I formed the aforementioned committee of colleagues from each library department to review the submissions. Using a similar ranking process to those for conference proposals or grant reviews, I assigned numerical value to the different form categories and brought individuals from across the library together to rank and discuss. This collaboration allowed us to use knowledge from different departments to evaluate the submissions, and to provide a sense of transparency in the decision making process. In the future, we plan to increase the perspectives by inviting faculty, students, and other community members to serve on

---

[6] Dorothy Berry, "Submission Form," Houghton Technical Services, last modified February 4, 2020, https://wiki.harvard.edu/confluence/display/HoughtonTechnicalServices/Submission+Form.

the committee.

I implemented this process at Houghton Library for a first run cycle in 2019, and for the first time, we were selecting projects through an equitable process that was open across the library. The digital projects committee reviewed 17 project proposals ranging from selections of three or four items to entire archival collections. This process required more work from submitters and committee members, but it also more accurately reflected the amount of work and planning that needs to go into successful digital projects. By making broader representation of the collections a weighted review factor, we required all submitters to think about how their proposal fits into our larger library goals around diversity and inclusion.

Ongoing workflows create stability and allow for long-term planning, but sometimes there are opportunities to divert resources (or receive new resources) to dedicate completely to projects that directly increase representation. In the 2021-22 fiscal year, we have paused the project submission process to focus on digitizing only materials related to African American history and culture for the "Slavery, Abolition, Emancipation, and Freedom: Primary Sources from Houghton Library" (SAEF) project.[7] The idea to focus entirely on SAEF while pausing other projects was complex. New, systematic workflows are essential in establishing best practices and inviting new voices, but it can be beneficial to create special projects that exist outside the proposal process that are specifically designed to increase access to underrepresented peoples' materials—especially when it is clear that there are large gaps. Our projects in progress were all fantastic, and many of them are in high demand. However, in the height of the summer of 2020, I received multiple requests from across the library for images of African Americans from our digital collections for social media or blog posts. I was struck both by how unfamiliar most of our library staff was with the holdings in the area, and also by how truly scant the available digitized options were. Moreover, there were purely practical reasons around a total pause. Limited resources may not be able to accomplish the ongoing project workflow alongside large special projects. Being flexible and open to possibilities is key in a constantly shifting cultural and institutional project context. Looking at our institutions' track records, special projects like SAEF are often the necessary response to long histories of exclusion.

An issue that comes up with this sort of work is the argument that it can't be done because there is no one on staff with particular knowledge about marginalized peoples' materials. This is one of my least favorite excuses for two reasons. First, as librarians and archivists we are perhaps the most adept folks at gaining familiarity with new material—every collection processed requires digging into a new person, history, and geography. Second, since none of us have unlimited staff resources, this argument either locks out all projects that don't focus on the topics on which current staff are experts, or it promotes hiring people from marginalized backgrounds into contingent

---

[7] Lydialyle Gibson, "Bringing Black History to Light," *Harvard Magazine*, July 15, 2020, https://www.harvardmagazine.com/2020/07/houghton-library-digitizes-african-american-history-materials.

positions that extract knowledge without providing long-term professional support.

While we can't all become true experts on every collecting area, recruiting colleagues with interests in relevant areas and committing to open communication and education can create a shared knowledge base that has benefits that reach outside any discrete project. For SAEF, I wanted to bring in colleagues from across special collections' departments: technical services, collections, and public services. I invited a public services colleague with a deep commitment to diversity and inclusion in libraries, a collections colleague from our modern books and manuscripts area (where most of the selected materials reside), and a new digital archivist colleague, who has been a great assistance to our thinking about digital collections.

Our project team has learned two major lessons from this pause for special projects: be willing to shift priorities and be willing to dream big and leap obstacles. We are all working with limited resources. Centering materials we have previously ignored requires reallocations but can lead to unexpected outcomes. SAEF has led to cataloging and identification of previously unknown collections materials, outreach and instruction to classes that previously had no prior relationship to the library, and new technological workflows that may be adopted campus-wide. Planning large initiatives is ambitious. When trying to increase representation, especially in response to community demands, there can be a desire for any project to be the end-all-be-all. Promising progress rather than perfection serves the difficult truth that problems that took years to create may take years to solve.

Developing a digital collections program from the ground up is a difficult journey with ups and downs. Integrating more equitable representation as a core value from the outset establishes a foundation for the type of library that our history requires and that our patrons demand. We've been able to move towards this goal by establishing clearly documented workflows, but also by being willing to shift gears when the moment is right.

The example of Houghton Library's digital collections program is far from the only way a library can increase its digital holdings while centering the margins. Archives and special collections have a rich history of speaking neutrality while making strong value judgments, and digital collections can provide one of the most democratized access points we have. We all have a responsibility to make sure our work is truly serving the greater good by fulfilling our goal to increase access to information and knowledge.

BIBLIOGRAPHY

"About Harvard Library." Harvard Library. Accessed September 1, 2021.
　　　https://library.harvard.edu/about/about-harvard-library.

Aston, Andy, Carolyn Caizzi, Kathleen Cameron, Jason Casden, Tim Clarke, Tom Cramer,
　　　Cristela Garcia-Spitz et al. *DLF Project Managers Toolkit*. DLF Wiki. Accessed
　　　August 24, 2021. https://wiki.diglib.org/DLF_Project_Managers_Toolkit.

Berry, Dorothy. "Submission Form." Houghton Technical Services. Last modified February
　　　4, 2020.
　　　https://wiki.harvard.edu/confluence/display/HoughtonTechnicalServices/Submis
　　　sion+Form.

Gibson, Lydialyle. "Bringing Black History to Light." *Harvard Magazine*, July 15, 2020.
　　　https://www.harvardmagazine.com/2020/07/houghton-library-digitizes-african-
　　　american-history-materials.

Jimerson, Randall C. "Archivists and Social Responsibility: A Response to Mark Greene."
　　　The American Archivist 76, no. 2 (Fall/Winter 2013): 335-45.
　　　http://www.jstor.org/stable/43490358.

Smith, Kathlin. "Mellon Foundation Funds CLIR's Digitizing Hidden Collections and
　　　Archives: Amplifying Unheard Voices." *CLIR News*, February 17, 2021.
　　　https://www.clir.org/2021/02/mellon-foundation-funds-amplifying-unheard-voi
　　　ces/.

Ziegler, Sophia. "Digitization Selection Criteria as Anti-Racist Action." *Code4Lib Journal*
　　　45 (August 9, 2019). https://journal.code4lib.org/articles/14667.