



The Lighting the Way Handbook

Case Studies, Guidelines, and Emergent Futures for Archival Discovery and Delivery



This project was made possible in part by the Institute of Museum and Library Services, through grant LG-35-19-0012-19. The views, findings, conclusions, or recommendations expressed in this publication do not necessarily represent those of the Institute of Museum and Library Services.

Edited by Mark A. Matienzo and Dinah Handel

STANFORD UNIVERSITY LIBRARIES | OCTOBER 2021



This work is licensed under a Creative Commons Attribution 4.0 International License.

Citation

Mark A. Matienzo and Dinah Handel, eds. 2021. *The Lighting the Way Handbook: Case Studies, Guidelines, and Emergent Futures for Archival Discovery and Delivery*. Stanford, CA: Stanford University Libraries. <https://doi.org/10.25740/gg453cv6438>.

Credits and Acknowledgements

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Unless otherwise noted, all authors contributed equally to the text of their chapter.

Funding

This project was made possible in part by the Institute of Museum and Library Services, through grant LG-35-19-0012-19. The views, findings, conclusions, or recommendations expressed in this publication do not necessarily represent those of the Institute of Museum and Library Services.

Project Team, Stanford University Libraries

Primary Project Team: Mark A. Matienzo, Assistant Director for Digital Strategy and Access and Project Director; Dinah Handel, Digitization Service Manager; Camille Villa, Digital Library Software Developer; Supavadee Kiattinant, Administrative Associate

Additional support from: Tom Cramer, Associate University Librarian and Director, Digital Library Systems and Services; Franz Kunst, Archivist; Sally DeBauche, Digital Archivist; Glynn Edwards, Assistant Director, Department of Special Collections; Josh Schneider, University Archivist

Working Meeting Facilitators

Hillel Arnold, Rockefeller Archive Center; Audra Eagle Yun, UC Irvine; Max Eckard, Bentley Historical Library, University of Michigan; Dinah Handel, Stanford University Libraries; Wendy Hagenmaier, Georgia Tech; Julie Hardesty, Indiana University; Linda Hocking, Litchfield Historical Society; Mark A. Matienzo, Stanford University Libraries; Gregory Wiedeman, University at Albany, SUNY

Participant Advisors

Amelia Abreu, UX Night School; Hillel Arnold, Rockefeller Archive Center; Elvia Arroyo-Ramírez, UC Irvine; Dorothy Berry, Harvard University; Audra Eagle Yun, UC Irvine; Max Eckard, Bentley Historical Library, University of Michigan; Amanda Ferrara, Seeley G. Mudd Manuscript Library, Princeton University; Geoff Froh, Densho; Julie Hardesty, Indiana University; Linda Hocking, Litchfield Historical Society; Sara Logue, Seeley G. Mudd Manuscript Library, Princeton University; Sandra Phoenix, HBCU Library Alliance; Greg Wiedeman, University at Albany, SUNY

Table of Contents

Introduction to <i>The Lighting the Way Handbook</i> <i>Mark A. Matienzo and Dinah Handel</i>	1
SECTION 1: CASE STUDIES	13
Connecting on Principles: Building and Uncovering Relationships through a New Archival Discovery System <i>Renee Pappous, Hannah Sistrunk, and Darren Young</i>	15
Access is People: How Investing in Digital Collections Labor Improves Archival Discovery & Delivery <i>Stephanie Becker, Anne Kumer, and Naomi Langer</i>	29
Facilitating Seamless Access Through Collaborative Workflows, Advocacy, and Communication <i>Martha Anderson, Max Eckard, Melanie Griffin, Emiko Hastings, Deb Kulczak, Chris Powell, Olga Virakhovskaya, Caitlin Wells, and Katrina Windon</i>	39
SECTION 2: ASSESSING AND APPLYING STANDARDS AND BEST PRACTICES	53
Lost Without Context: Representing Relationships between Archival Materials in the Digital Environment <i>Jodi Allison-Bunnell, Maureen Cresci Callahan, Gretchen Gueguen, John Kunze, Krystyna K. Matusiak, and Gregory Wiedeman</i>	55
Maximizing Good: An Inquiry-Based Approach to Minimal Description for Online Archives <i>Sarah Dorpinghaus, Cory Lampert, Rebecca Pattillo, and Kyna Herzinger</i>	73
Playing to our Strengths: Self-Assessment Criteria for Access and Discovery in Small Archives <i>Stefana Breitwieser, Amanda Demeter, Sophie Glidden-Lyon, Amanda Murray, Lori Myers- Steele, and Kate Philipson</i>	85
SECTION 3: EMERGENT OPPORTUNITIES	97
The Power of Parallel Description: Wikidata and Archival Discovery <i>Kelli Babcock, Regine Heberlein, Anna Björnsson McCormick, Elizabeth Russey Roke, Greta Kuriger Suiter, and Ruth Kitchin Tillman</i>	99
Inviting and Honoring User-contributed Content <i>Katherine Crowe, Katrina Fenlon, Hannah Frisch, Diana Marsh, and Victoria Van Hying</i>	115
A Call to Action: User Experience & Inclusive Description <i>Faith Charlton, Christa Cleeton, Alison Clemens, Betts Coup, Zoë Hill, and Jessica Tai</i>	133
Speeding Towards Remote Access: Developing Shared Recommendations for Virtual Reading Rooms <i>Elvia Arroyo-Ramírez, Annalise Berdini, Shelly Black, Greg Cram, Kathryn Gronsbell, Nick Krabbenhoft, Kate Lynch, Genevieve Preston, and Heather Smedberg</i>	141

Introduction to *The Lighting the Way Handbook*

Mark A. Matienzo and Dinah Handel

Abstract: *The Lighting the Way Handbook: Case Studies, Guidelines, and Emergent Futures for Archival Discovery and Delivery* represents the synthesis of work undertaken by participants in the Lighting the Way Working Meeting, held virtually in April-May 2021. The Working Meeting was organized as a practitioner-focused strategic thinking opportunity intended to explore topics related to archival discovery and delivery. Working Meeting participants met in several facilitated sessions using techniques from Liberating Structures, an inclusive facilitation methodology, and followed an intentional progression of steps to generate and structure their ideas. This introduction contextualizes the work and identifies several themes across the submissions, as well as provides recommendations for future areas of work and considerations to programmatically support strategic work to improve archival discovery and delivery. Our recommendations to sustain this work include 1) establishing an investment in understanding collaborative models, power relations, and organizational positioning of this work; 2) ensuring time and space for strategic planning and advocating using care-focused methods; and 3) identifying ways in which to create and sustain communities of practice.

Background

The Lighting the Way Handbook: Case Studies, Guidelines, and Emergent Futures for Archival Discovery and Delivery is the culmination of the Lighting the Way Working Meeting, a series of virtual workshops held in April-May 2021 as part of the Lighting the Way project. The Working Meeting built on the previous efforts of the project, including the February 2020 Forum (Matienzo et al. 2020), an in-person event with 70 participants focusing on facilitated discussion. While the Forum provided a starting point, the Working Meeting was intended as an opportunity for selected participants to investigate a topic related to future-oriented opportunities to improve archival discovery and delivery (as defined below), including case studies or identification of strategic opportunities in greater depth. Participants subsequently wrote these in-depth explorations into the submissions included in this publication.

What is archival discovery and delivery?

Archival discovery and delivery is the phrase used by the Lighting the Way project describing what people, processes, and systems do to support finding, accessing, and using material from archives and special collections. While the project initially focused on integration between systems as its primary area of analysis, early project investigations and the discussions at the Forum led us to realize that this work is necessarily performed by people in a variety of roles – not just archives workers, but library workers, technology workers, and others with varying skill sets, areas of expertise, levels of responsibility, and positional power within their institutions. Part of the broader challenge is to determine how to effectively align the people, processes, and systems that fit into this broader function. It requires close collaboration across job roles and responsibilities, departments, and institutions, like

other areas of work, but in some senses is the least understood given these complexities. “Archival discovery and delivery” is thus intended to underscore the complexity and interdependence of the work, and to take a more expansive view of this work than one focused solely on archival functions as currently understood or as technical development and implementation completed and supported by IT service providers with archives workers as “clients” (e.g. as described by Shaw, Adler, and Dooley 2017). Following Weber (2017), the project viewed archival discovery and delivery as relying on an “ecosystem of systems” in a wide variety of functional roles.

The Working Meeting: selection process, structure, and facilitation

Participants for the Working Meeting were solicited through an open call for participants, distributed widely to professional communities and email lists. Attendance at the Forum was not a prerequisite, and Forum attendance also did not guarantee that applications would be selected. While open to individuals, the call for participants encouraged submissions for groups of 3-6 individuals and asked for a brief abstract of the potential topic intended for exploration. The call for participants received 24 submissions with a total of 100 individuals, and the project team selected nine groups to participate in the Working Meeting. Two additional groups were encouraged to consider merging with a selected group and agreed to do so given the similarity in their topics. Each of the nine groups were assigned a designated facilitator. In addition to the nine selected groups, the project team invited a tenth group of participants to provide a written contribution given limitations on the number of available facilitators for the Working Meeting.

The Working Meeting was organized as four two-hour sessions held on Zoom; the first and last sessions were plenary sessions with breakouts, and the second and third sessions were scheduled separately for each group and their facilitator. Each session used a variety of facilitation methods drawn from Liberating Structures (Lipmanowicz and McCandless 2014; n.d.), a framework of facilitation techniques intended for participatory and inclusive events, as well as activities drawn from its associated community of practice and additional sources. The structure of the sessions was further informed by the adoption of *strategy knotworking* (McCandless and Schartau 2018), a refinement and application of Liberating Structures that applies its methods and structure to inform strategic planning through an iterative exploration of six questions:

1. **Purpose:** what is the fundamental justification for the existence of our work?
2. **Context:** what is happening around us that demands creative change?
3. **Challenge:** what paradoxical challenges must we face to make progress?
4. **Baseline:** where are we starting, honestly?
5. **Ambition:** given our purpose, what seems possible now?
6. **Action and evaluation:** how are we acting our way toward the future, evaluating what is possible as we go?

Each session focused one or two of these six questions through the use of the facilitated activities, and facilitators were given some discretion to adapt sessions as needed for each group.¹ Participant groups were encouraged by the facilitators and project team to address and incorporate elements of the six questions of strategy knotworking in some manner within their chapters for inclusion in *The Lighting the Way Handbook*, although groups did not otherwise have specific structural requirements to follow. Throughout the Working Meeting sessions, participants self-organized to complete the work on their submissions published within this volume.

Organization and contents

The Lighting the Way Handbook's chapters are organized into three primary sections. While any organizational scheme is reductive and may gloss over the nuanced arguments within each chapter, the groupings are intended to reflect the similarities in structure or focus for the chapters. The editors of this volume hope that the organization allows readers to see the resonances across the varying chapters and helps them respectively amplify the arguments or positions included in each. In addition, while each group worked on their chapters independently, there are clear thematic connections across some of them. These are identified and discussed further in the *Emerging themes* section that follows this one.

The first section, **Case Studies**, comprises case studies of specific work related to archival discovery and delivery (understood in a broad sense) within specific institutions. While each case study has an institutional focus and acknowledges completed work or the current state of affairs within each institution, they also acknowledge future work to come, or areas for broader consideration related to archival discovery and delivery. The section begins with “Connecting on Principles: Building and Uncovering Relationships through a New Archival Discovery System” by Renee Pappous, Hannah Sistrunk, and Darren Young, which focuses on the development and implementation of the Rockefeller Archive Center’s latest iteration of its front-end system supporting archival discovery and delivery. It explores collaborative work, informed by descriptive standards, and the change in working relationships necessary to complete and sustain the system and work that supports it. The second chapter, “Access is People: How Investing in Digital Collections Labor Improves Archival Discovery and Delivery” by Stephanie Becker, Anne Kumer, and Naomi Langer, investigates the necessity of valuing the labor of staff responsible for creating and stewarding digital collections. The chapter is situated in the context of a case study of the authors’ experience creating and serving on a Digitization Governance Committee at Case Western Reserve University and includes tangible guidance on what this looks like in practice. The final chapter in this section is Martha Anderson, Max Eckard, Melanie Griffin, Emiko Hastings, Deb Kulczak, Chris Powell, Olga Virakhovskaya, Caitlin Wells, and Katrina Windon’s “Facilitating Seamless Access Through Collaborative Workflows, Advocacy, and Communication.” This chapter focuses on the complexities and necessity of collaboration within a given institution to support effective archival discovery and delivery. It investigates two specific case studies of collaborative cultures at the University of Michigan and University of Arkansas, the complex ecosystems of systems within each, and identifies generalizable approaches for fostering ongoing collaboration.

¹ For a detailed discussion of the activities held within the Working Meeting, see the Playbook appendix of the Lighting the Way project’s final report (Handel and Matienzo 2021).

The second section, **Assessing and Applying Standards and Best Practices**, focuses on chapters that engage specifically with standards and best practices that impact archival discovery and delivery. The first chapter within this section, “Lost Without Context: Representing Relationships Between Archival Materials in the Digital Environment” by Jodi Allison-Bunnell, Maureen Cresci Callahan, Gretchen Gueguen, John Kunze, and Krystyna K. Matusiak, argues for the importance of expressing archival context in systems supporting archival discovery and delivery. The chapter identifies a set of principles for the design of archival discovery and delivery principles informed in part by leveraging existing standards for archival description, like *Describing Archives: A Content Standard* and Encoded Archival Description. The following chapter is Sarah Dorpinghaus, Cory Lampert, Rebecca Pattillo, and Kyna Herzinger’s “Maximizing Good: An Inquiry-Based Approach to Minimal Description for Online Archives,” which considers the impact of minimal processing and descriptive practices, such as those described as part of the “More Product, Less Process” approach described by Greene and Meissner (2005). The authors analyze assumptions about the impact of minimal description on digital collections and provide recommendations to realign archival practice in a systematic manner, including usability, supporting systems, labor issues, and more. The chapter “Playing to our Strengths: Self-Assessment Criteria for Access and Discovery in Small Archives,” by Stefana Breitwieser, Amanda Demeter, Sophie Glidden-Lyon, Amanda Murray, Lori Myers-Steele, and Kate Philipson concludes this section. The authors focus on specific challenges experienced by and strengths available to small archives and the subsequent understanding of what successful archival discovery and delivery looks like for these programs. It also provides a guided set of questions for self-assessment for workers within small archives to help audit practices and define alternative and sustainable visions of success for them.

The final section in *The Lighting the Way Handbook*, **Emergent Opportunities**, contains chapters that focus specifically on exploring new opportunities. While each chapter within this section acknowledges and leverages past work related to archival discovery and delivery, they also advocate for more exhaustive and programmatic work in their areas of focus. These chapters also strongly advocate for situating this work in relation to community engagement and development in relation to both professional communities of practice and community-supported and -led efforts outside of archives, library, and technology spheres. The section begins with “The Power of Parallel Description: Wikidata and Archival Discovery” by Kelli Babcock, Regine Heberlein, Anna Björnsson McCormick, Elizabeth Russey Roke, Greta Kuriger Suiter, and Ruth Kitchin Tillman, which advocates for the use of Wikidata in archival descriptive workflows. The authors investigate this as a divergence from existing archival practice and provide a set of actionable recommendations for how archivists can begin working with Wikidata. Katherine Crowe, Katrina Fenlon, Hannah Frisch, Diana Marsh, and Victoria Van Hyning’s chapter, “Inviting and Honoring User-contributed Content,” investigates the challenges to and potential impact of integrating user-contributed content to the landscape of data managed by libraries, archives, and museums. The authors review potential models for including user-contributed content, unpack its relationship to supporting Indigenous collections and community collaborations and generative research practices, and articulate its ongoing value and the responsibilities of library, archives, and museum workers to incorporate it into their descriptive ecosystems. The third chapter, Faith Charlton, Christa Cleeton, Alison Clemens, Betts Coup, Zoë Hill, and Jessica Tai’s “A Call to Action: User Experience and Inclusive Description,” focuses on leveraging user experience research and design to allow archivists to understand the impact of reparative and inclusive description projects. The chapter concludes with a set of recommendations on how individuals, institutions, and professions can center users in descriptive practice, including through advocating for the creation of a professional community of practice focused on usability within archives. The section concludes with “Speeding

Towards Remote Access: Developing Shared Recommendations for Virtual Reading Rooms” by Elvia Arroyo-Ramírez, Annalise Berdini, Shelly Black, Greg Cram, Kathryn Gronsbell, Nick Krabbenhoeft, Kate Lynch, Genevieve Preston, and Heather Smedberg. Through their experience with developing mediated delivery systems for digital archives at seven institutions, the authors focus on the creation of a shared framework to define requirements and considerations for building comparable systems, with analysis into needs including advocacy and outreach, resources, users and use cases, ethical concerns, copyright, and system interoperability. They incorporate recommendations throughout their analysis, and advocate for inter-institutional partnerships and the development of a professional community of practice to support this burgeoning form of ecosystem supporting archival discovery and delivery.

Emerging themes within *The Lighting the Way Handbook*

While groups participating in the Working Meeting were part of a larger cohort focused on investigating topics related to archival discovery and delivery, they undertook their work primarily independently from one another. Accordingly, the project team wanted to identify, connect, and synthesize the themes across the chapters within this publication. Similarly, these threads also connect to broader points and conversations that also surfaced during Working Meeting activities and the entire project. This synthesis also serves in part as a starting point for the recommendations and opportunities described in the following section.

The ecosystem of systems in archival discovery and delivery

While the project purposely identified a wide range of systems supporting archival discovery and delivery, the realities of system integration mean that the ecosystem of systems can be complex. For instance, in their chapter, Arroyo-Ramírez et al. emphasize the importance of integration and interoperability of virtual reading rooms as ecosystems supporting mediated access to archives. Despite significant advancements in the archives and library sectors, certain kinds of integration continue to remain challenging. The case studies in the chapter by Anderson et al. describe the challenges when organizational units use parallel instances of similar systems, while also emphasizing the importance of integration supported by application programming interfaces and automated communication similarly in the way that Pappous et al. do within their chapter. Allison-Bunnell et al. and Dorpinghaus et al. both explore the impact on users when there are limits to the effective interchange of data between systems, such as between a collection management system and discovery and delivery supported by digital library systems. The chapter by Crowe et al. highlights that despite advances in archival systems, there are nonetheless significant gaps in the ability of those systems to incorporate user-contributed content. **As the complexity of collections, systems, and user requirements continue to evolve, the complexity of the ecosystem of systems will also evolve and grow.**

The COVID-19 pandemic and archival discovery and delivery

The COVID-19 pandemic has had a broad impact on many areas of research (“Beyond the Pandemic” 2021). **Subsequently, the archives and library sectors have similarly considered how to adjust their operations and undertake new initiatives to better support remote research.** Unsurprisingly, several authors within *The Lighting the Way Handbook* describe the motivations and impacts within their chapters. Pappous et al. describe the pandemic's effects on tangible work such as user testing of a newly developed archival discovery system as well as an understanding of longer-term changes to

how research needs to be supported. Arroyo-Ramírez et al. and Becker et al. also acknowledge how the pandemic has led to stakeholder needs evolving and demanding nearer term changes requiring substantial investment. The pandemic, of course, is also a global public health crisis with substantial impact on marginalized communities. Crowe et al. acknowledge this impact on Indigenous communities in particular, and describe how taking user-contributed content seriously as part of pandemic response would have a potentially positive impact on equity for collections access and use.

Resource-sensitive operations, valuing labor, and impact on strategy and advocacy

The overwhelming majority of chapters within this volume argue for **improving operations and taking strategic directions that will be sustainable and sensitive to the realities of resourcing within a given institutional or community context**. Tools to assess capacity, such as the framework provided by Breitwieser et al. for self-assessment by small archives, are proving themselves to be central to informing what archives programs can not only support, but what they can aspire to undertake. Similarly, other contributors, like Dorpinghaus et al., Becker et al., and Arroyo-Ramírez et al., have looked to the recently published Total Cost of Stewardship framework (Weber et al. 2021) to understand both capacity for and operational impact of projects and initiatives related to archival discovery and delivery.

The contributors also recognize that **supporting both operations and new strategic investments requires substantial advocacy** to ensure that resource allocators understand not only archival discovery and delivery, but the broader areas of archives, library, and technology work essential to these programs. For instance, Pappous et al. recognize that sustainability in the case of supporting the latest iteration of the Rockefeller Archive Center's DIMES discovery system requires investment of staff in its success and maintenance and may require changes in working relationships. The framework provided by Arroyo-Ramírez et al. includes an analysis of how to provide programmatic support for the development and sustenance of virtual reading rooms, including ensuring administrators and resource allocators understand the motivation and necessary resources.

Several chapters, including Becker et al., Arroyo-Ramírez et al., Dorpinghaus et al., and Breitwieser et al., explicitly acknowledge that **valuing the labor necessary to undertake the kinds of strategic opportunities described therein is important to their success**. Relatedly, these authors also note the negative impact that the use of temporary positions and grant funding can have on this work. Unfortunately, recent research by Blumenthal et al. (2020) and past work such as the Levy Report (Levy and Robles 1984) suggests that **the level of advocacy necessary means that resource allocators still do not fully understand the work to support archival discovery and delivery**. It also reflects a multigenerational issue in a broad sense that impacts archives and special collections. Arroyo-Ramírez et al. and Becker et al. indicate that Blumenthal et al.'s research underscores that this systemic misunderstanding is a continued threat to making transformational strategic change to improve archival discovery and delivery. Despite recent work analyzing the impact of grant-funded positions and providing ethical guidance for using term labor (Rodriguez et al. 2019; Tillman and Rodriguez 2020), term labor surfaces as not only a potential cause of this under-resourcing and misunderstanding, but also a symptom thereof. The reality is that these forms of austerity are deeply rooted both within archival programs as well as the larger institutions in which they exist (Rizzo 2021).

Rethinking the structure and practice of collaboration and organizational positioning

As stated above, work supporting archival discovery and delivery depends on a great deal of collaboration across a variety of roles and departments, and many of the chapters included in *The Lighting the Way Handbook* acknowledge this. **The institutional realities of working relationships across departments can have a significant impact on the efficacy of archival discovery and delivery.** This has been described in the three case studies by Pappous et al., Becker et al., and Anderson et al. In addition, the chapters by Babcock et al., Crowe et al., and Charlton et al. also acknowledge the importance of conscientious investment into collaboration with broader communities of users and stakeholders. Charlton, et al. also acknowledge that there needs to be “institutional responsibility” to make a commitment to user-centered practices to assess the impact of reparative and inclusive description, which includes commitments to structure the work equitably. The chapters by Breitwieser et al. and Arroyo-Ramírez et al. also describe the complexities of how archives workers need to think more broadly about collaboration in two quite different cases: supporting archival discovery and delivery for small archives programs, and for development of virtual reading rooms. While these two specific contexts vary significantly, they both nonetheless recognize the need for broader internal partnerships to allow strategic initiatives and operational needs to come to fruition and to be sustained. Weber (2017) notes there have been significant changes in the last 15 years in how archives and special collections are organized and positioned within research libraries. **Unsurprisingly, this organizational positioning also impacts collaboration, and as she suggests, there may be further affordances to refining organizational structures to improve archival discovery and delivery.**

Connections and tensions between standards, best practices, and stakeholder needs

Several chapters explore how **archival standards and best practices can both sustain and impede effective work on archival discovery and delivery.** While several contributions referenced *Describing Archives: A Content Standard* (most notably through its revised principles), the chapters by Allison-Bunnell et al. and Pappous et al. are informed by that standard albeit in distinct ways. Allison-Bunnell et al.’s chapter focuses specifically on the importance of existing descriptive standards and foundational professional concepts, namely context, as essential to informing how archival discovery and delivery systems should function. In contrast, Pappous et al. (as well as Charlton et al. and Breitwieser et al., albeit to a lesser extent) use the revised DACS principles as a generative starting point to inform their implementation of a new version of their archival discovery system as well as the collaboration necessary to sustain it.

Several chapters also engage with the **tensions between the suitability of archival practices and standards in relation to addressing stakeholder needs.** Dorpinghaus et al. contest the assumed benefits of the “More Product, Less Process” method for minimal descriptive practices in assessing its impact on online access to digital archival collections. Breitwieser et al. argue that archives workers in small archives programs should advocate for best practices that appropriately serve their institutional context. Crowe et al. and Babcock et al. identify the challenges in applying thinking based upon existing standards to cases around better supporting user-contributed content and Wikidata respectively, and revisit how archives workers can step out of a role of authority and think differently about collaboration when working with each. Furthermore, chapters such as those by Arroyo-Ramírez et al., Crowe et al.,

and Charlton et al., also acknowledge areas in which professional communities of archives workers and their stakeholder communities have opportunities to further define best practices to support their respective areas of investigation. In some cases, contributors have identified the importance of creating communities of practice for their areas to support the process of defining best practices. Despite a history of user studies in archival research, Charlton et al. identify a lack of a community of practice supporting usability work in archives in a holistic manner. Arroyo-Ramírez et al. describe the importance of sharing knowledge across professional groups as central to the development of virtual reading rooms given their recent emergence.

Considering the path forward: recommendations and opportunities

The chapters within *The Lighting the Way Handbook* offer a variety of starting points in recognition that much of the work to improve archival discovery and delivery is still in formative stages. As such, additional investment of time, resources, and careful planning and exploration are necessary to undertake these areas of effort. This section identifies a set of potential recommendations and areas of consideration in service of the broader need to improve archival discovery and delivery. While not comprehensive, these areas for exploration provide a starting point for archives, library, and technology workers to explore and act upon strategic initiatives related to archival discovery and delivery and beyond.

Collaboration is essential and is impacted by both power relationships and the cultural norms between collaborating parties. Emerging needs for collaboration must further impact the organizational positioning for this work. As the Working Meeting and the subsequent chapters written by participants demonstrate, effective archival discovery and delivery, as well as transformational change, requires deep collaboration. Participants at the Forum and Working Meeting both recognized that they often faced structural challenges to collaboration. The discussion of the impact on resource allocation on tactical and strategic progress within these chapters addresses a disconnect between senior leadership and administration and the day-to-day lived realities of workers responsible for improving archival discovery and delivery. Many of the authors elucidated the importance of making work more transparent and have advocated for shared responsibility, but challenges remain to have this work be understood and resourced appropriately. Participants provided positive feedback on the facilitation methods to help identify, understand, and potentially address issues around these power relationships, but doing so with senior leadership and administration requires a substantial amount of trust. Participants also identified the wide variety of roles that need to be engaged for programmatic work around archives and delivery, such as Arroyo-Ramírez et al. have done for virtual reading rooms. Accordingly, there is an ongoing need to engage colleagues with relevant expertise and knowledge across these initiatives. However, as Shaw, Adler and Dooley (2017) describe, collaborating with technology staff in particular can be particularly challenging given cultural differences across units. While activities may likely be constrained because of these realities, there are nonetheless opportunities to improve and identify new models to undertake shared responsibilities for these programs (O'Meara 2013). This also may mean creating new governance structures for work such as those identified by Becker et al., or changing the organizational positioning of the work to ensure that it is supported adequately, such as the creation of new service teams or units charged with this responsibility.

Strategic planning for archival programs is essential, and care-focused and generative methods such as those used within the Lighting the Way project are of great value to practitioners. Accordingly, archives, library, and technology practitioners must undertake responsibility to apply these methods within their own institutions. Throughout the project, participants were highly appreciative to have the time and space to explore challenges and opportunities in relation to archival discovery and delivery. Participants in the Working Meeting were grateful to have the time for this exploration amidst the pandemic. However, participants also noted how difficult it is to set aside time to undertake similar work within their own institutions. While the project did not explore this gap in depth, there are clear questions about why such efforts have not been prioritized. However, as Blumenthal et al. (2020) describe, lack of strategic thinking and relevant advocacy will have a generational impact that leads to this work being inadequately supported over the longer term. Given the mostly positive experiences, participants noted a desire to learn more about the generative facilitation methods drawn from Liberating Structures and other sources. While many participants have inquired whether the project will be seeking additional funding to continue the project's model, the project team recognizes that it is fundamentally unsustainable for one group to be solely responsible for organizing these conversations even within the focused area of archival discovery and delivery. The project team is investigating the development of a broader set of concrete recommendations of how archival programs can use such methods to inform and structure the work they do, especially given that the methods can also be applied to smaller scale meetings outside of community convenings such as those within this project. The facilitation model used within the project has also been deeply informed by models of shared affective responsibility for archives (Caswell and Cifor 2016; O'Neill et al. 2017; Arroyo-Ramírez and Jones 2018), and similarly, it is up to practitioners to advocate for and incorporate a care-centered approach to this work. This itself can be supported by investments in using care-centered facilitation methods in strategic planning.

Supporting archival discovery and delivery requires creation of and participation of new communities of practice, as well as alignment with existing ones. Sustainability of communities of practice is a continued area of concern for archives, library, and technology practitioners, so the creation of new communities should be highly focused. A community of practice is a group with shared or common interest in an area of technical knowledge or professional activity. As described above, several chapters identified a clear need for communities of practice to support some of the more emergent areas of archival discovery and delivery, such as virtual reading rooms and usability and user experience. Participants throughout the project also noted the wide variety of communities of practice that already exist. These include nearly 50 sections of the Society of American Archivists (2019), the BitCurator Users Forum, the Digital Library Federation's Born Digital Access Working Group, and the many communities of practice that exist around specific software platforms or tools. Participants acknowledged that undertaking new or generative work in some communities was more challenging than others, but many of these communities of practice have a specific focus or mandate. There is a clear need around strategic and operational planning in support of not only archival discovery and delivery, but technology projects for archives more broadly. In any event, communities of practice can be challenging to sustain, so to best capitalize on these opportunities, it is essential to define a clear focus, mandate, and relationship to other communities.

Conclusion

The Lighting the Way Handbook, like the rest of the Lighting the Way project, is intended as a starting point to reflect both the current state of completed work, the changing relationship to standards and best practices, and emergent areas for further focused effort. The chapters within this publication represent the engagement of the participants in the practice of strategy knotworking, a specific application of the Liberating Structures facilitation framework focused on strategic planning. The focus of these contributions and the project is on archival discovery and delivery as an emergent and broader understanding of the work and systems needed to support effective access and use of archives. Through this process and the work of the project more broadly, the project team and participants recognized emerging themes related to the ecosystem of systems supporting archival discovery and delivery; the impact of the COVID-19 pandemic on both operations and planning; the connections to emerging models of resource-sensitive operations; and the connections and tensions between this work and professional standards and best practices. Our recommendations to sustain this work include establishing an investment in understanding collaborative models, power relations, and organizational positioning of this work; ensuring time and space for strategic planning and advocating for care-focused methods; and identifying ways in which to create and sustain communities of practice. While this requires a considerable investment from practitioners, the collective experience of participants and facilitators in the project demonstrate that this is essential to ensure the continued success of this work.

Acknowledgements

The authors of this introduction wish to acknowledge the contributions of all the participants in the Lighting the Way Working Meeting and contributors to this publication, without whom it would not exist. In addition, we would like to acknowledge the work of the Working Meeting's additional facilitators – Hillel Arnold, Audra Eagle Yun, Max Eckard, Wendy Hagenmaier, Julie Hardesty, Linda Hocking, Greg Wiedeman – for their willingness to commit the time and energy to organize, contribute, and lead a virtual event during a challenging time. Their efforts and dedication ensured that the Working Meeting were inclusive, generative, and well-organized for the participants. We also want to thank the additional Participant Advisors on the project, including Amelia Abreu, Dorothy Berry, Amanda Ferrara, Geoff Froh, Sara Logue, and Sandra Phoenix, for their continued service and thoughtful feedback. Combined, our group of project participants, facilitators, and advisors provided an essential role in helping shape our project and informing its future directions.

References

- Arroyo-Ramírez, Elvia, and Jasmine Jones. 2018. "Applying Radical Empathy Framework in Archival Practice." Presented at the Society of California Archivists Webinar, September 17.
- "Beyond the Pandemic: The Future of the Research Enterprise in Academic Year 2021-22 and Beyond." 2021. Coalition for Networked Information. <https://www.cni.org/wp-content/uploads/2021/08/report-CNI-X-RT-Research-June-2021-FINAL.pdf>.
- Blumenthal, Karl, Peggy Griesinger, Julia Y. Kim, Shira Peltzman, and Vicky Steeves. 2020. "What's Wrong with Digital Stewardship: Evaluating the Organization of Digital Preservation Programs from Practitioners' Perspectives." *Journal of Contemporary Archival Studies* 7, Article 13. <https://elischolar.library.yale.edu/jcas/vol7/iss1/13/>.
- Caswell, Michelle, and Marika Cifor. 2016. "From Human Rights to Feminist Ethics: Radical Empathy in Archives." *Archivaria* 81.
- Greene, Mark, and Dennis Meissner. 2005. "More Product, Less Process: Revamping Traditional Archival Processing." *American Archivist* 68 (2): 208–63.
- Handel, Dinah, and Mark A. Matienzo. 2021 (forthcoming). "Facilitating and Illuminating Emergent Futures for Archival Discovery and Delivery: The Final Report of the Lighting the Way Project." Stanford, CA: Stanford University Libraries. <https://doi.org/10.25740/jm302fq5311>.
- Levy, Sidney J., and Albert G. Robles. 1984. "The Image of Archivists: Resource Allocators' Perceptions." Study #722/01. Chicago: Social Research, Inc. <https://catalog.hathitrust.org/Record/004031185>.
- Lipmanowicz, Henri, and Keith McCandless. 2014. *The Surprising Power of Liberating Structures: Simple Rules to Unleash A Culture of Innovation*. 1st edition. S.l.: Liberating Structures Press.
- . n.d. "Liberating Structures - Introduction." Accessed August 9, 2021. <https://www.liberatingstructures.com/>.
- Matienzo, Mark A., Dinah Handel, Josh Schneider, and Camille Villa. 2020. "Lighting the Way: A Preliminary Report on the National Forum on Archival Discovery and Delivery." Stanford University Libraries. <https://purl.stanford.edu/rt078dm2196>.
- McCandless, Keith, and Johannes Schartau. 2018. "Liberating Strategy: Surprise and Serendipity Put to Work." <https://liberatingstructures.eu/wp-content/uploads/2018/12/Liberating-Strategy.pdf>.
- O'Meara, Erin. 2013. "No One Cooks the Bacon Alone." Presented at the Past Forward: Meeting Stakeholder Needs in 21st Century Special Collections, New Haven, CT, June 4. <https://www.youtube.com/watch?v=xnQ5G1fH5-8>.

- O'Neill, Shannon, Elvia Arroyo-Ramírez, Molly Brown, Dinah Handel, Jasmine Jones, Rachel Mattson, Giordana Mecagni, Holly A Smith, and Kelly Wooten. 2017. "Radical Empathy in Archival Practice." Presented at the Archives 2017: Annual Meeting of the Society of American Archivists, Portland, Oregon, USA, July 28. <https://archives2017.sched.com/event/ABGy/301-radical-empathy-in-archival-practice>.
- Rizzo, Caitlin. 2021. "Archives After Austerity." *Acid Free*, 2021. <http://laacollective.org/work/archives-after-austerity>.
- Rodriguez, Sandy, Ruth Tillman, Amy Wickner, Stacie Williams, and Emily Drabinski. 2019. "Collective Responsibility: Seeking Equity for Contingent Labor in Libraries, Archives, and Museums." OSF. <https://osf.io/af9hz/>.
- Shaw, Seth, Richard C. Adler, and Jackie M. Dooley. 2017. *Demystifying IT: A Framework for Shared Understanding between Archivists and IT Professionals*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/C3192N>.
- Society of American Archivists. 2019. "SAA Sections." Society of American Archivists. October 23, 2019. <https://www2.archivists.org/groups/sections>.
- Tillman, Ruth Kitchin, and Sandy Rodriguez. 2020. "Collective Equity: A Handbook for Designing and Evaluating Grant-Funded Positions." <https://doi.org/10.26207/6p4a-md61>.
- Weber, Chela Scott. 2017. *Research and Learning Agenda for Archives, Special, and Distinctive Collections in Research Libraries*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/C3C34F>.
- Weber, Chela Scott, Martha Conway, Nicholas Martin, Gioia Stevens, and Brigitte Kamsler. 2021. *Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/ZBH0-A044>.

Section 1: Case Studies

Connecting on Principles: Building and Uncovering Relationships through a New Archival Discovery System

Renee Pappous, Hannah Sistrunk, and Darren Young

Abstract: The Rockefeller Archive Center has recently developed and released DIMES, a new front-end system for archival discovery and delivery, along with the infrastructure that integrates DIMES with systems for archival data and request management. In this case study, a team of DIMES contributors will outline why RAC archivists chose to design a new discovery system that supports the Describing Archives: A Content Standard Statement of Principles, and how our collecting areas — namely, the records of major philanthropies and the papers of the Rockefeller family — are uniquely primed for this type of discovery. We will then detail and evaluate aspects of how we built DIMES, emphasizing collaborative work involving contributors across the organization, including data cleanup and enhancement projects, usability testing, participatory design activities, and a rollout program in support of an internal launch of the site. Finally, we will discuss the projects' future, including ongoing maintenance and user-centered development work. With this model, we hope to demonstrate how archival institutions can harness the relationships found amongst their staff and their archival data to create and manage the transition to sustainable, meaningful systems that benefit users.

Introduction

In February 2021, the Rockefeller Archive Center (RAC) launched DIMES (2021b), a new system for archival discovery and delivery. The development of DIMES brought together the perspectives, expertise, and labor of staff from all program areas at the RAC for the purpose of providing flexible, ethical, and equitable access to the institution's collections. Using DIMES as a case study, three archivists representing the RAC's Processing, Reference, and Digital Strategies program teams will model how human-centered discovery can build and uncover relationships in archival discovery systems and between people in the ways we build and implement these systems.

About the Rockefeller Archive Center

The RAC is a repository of historical materials and a research center dedicated to the study of philanthropy. It holds the archives of major foundations, cultural organizations, research institutions, and many individuals associated with these organizations. Many of the donor organizations are currently active and transfer records to the archives on an ongoing basis. The RAC is organized into four archival program areas consisting of about 30 staff members: Reference, Processing, Collections Management, and Digital Strategies. Additionally, there is a Research and Education program and an IT team of two.

What is DIMES?

DIMES is the culmination of Project Electron (RAC 2021e), a RAC initiative to provide broad and equitable access to digital records through building sustainable, user-centered, and standards-compliant infrastructure scoped for acquiring, managing, and preserving records. It replaces the previous discovery system released in 2012, also called DIMES, which was based on the eXtensible Text Framework (XTF). This system was a success in that it made our archival data publicly available and searchable online. However, as our reference numbers increased and our collections continued to grow, there was room for improvement in the system's performance, interoperability, and accessibility.

In particular, the system's reliance on a traditional archival finding aid presentation of information was limited in its ability to allow users to browse the contextual relationships that bring meaning to our archival data. Researchers could find information relevant to their query in the search results, but those results could only take them to the associated finding aid and would not show the larger context that could enrich their understanding of the information and provide new avenues for their research, such as a shared creator.

Additionally, the XTF framework that underpinned the previous system relied on EAD XML documents representing entire collection-level resource records in ArchivesSpace, which increased the overhead of processing updates. Any change made within ArchivesSpace to an object within a resource record such as a note, date, or file title required ArchivesSpace to export all of the data within that resource record in order for the system to process and display the change in the public-facing description.

The emphasis on relationships within the updated version of DIMES can be observed within the very architecture of the system (Figure 1). DIMES encompasses not only the public-facing website where users can search and access archival materials but also the infrastructure to fetch, merge, transform, and index data (RAC 2021d, 2021h), an image pipeline (RAC 2021c), and an IIIF image server. While ArchivesSpace is our primary data source, the data pipeline (RAC 2021a) can accept other sources, like Wikidata, that make their data available through an application programming interface (API). The data pipeline moves changes made in ArchivesSpace to the frontend discovery system more efficiently than the previous version of DIMES because it employs a small group of automated services to prepare and transform the data pushed from the ArchivesSpace API. The discrete functions of these services and their relationships to one another enable exports of smaller units of data than the entire resource records pushed through by the old system and improve our ability to find where errors have occurred in the pipeline when something goes wrong.

Before the data is pushed to DIMES, the data pipeline transforms it to comply with the Rockefeller Archive Center data model (2021g). This data model liberates archival data from a strictly document-based presentation and places it within relationships that connect records, people, and activities. One way that researchers can now interact with those relationships when using DIMES is through the browsing that "arrangement maps" (RAC 2021f) support. Arrangement maps are tree representations the RAC creates for all collections, sub-collections, and sub-components (series, subseries, etc.) that originate from a shared creator. They extend the structures available in ArchivesSpace to handle accruals to existing collections that exist as separate resource records. The significance of the relationships the arrangement maps reveal in the archival collections of the RAC will be further explored in the next section.

Finally, once data is in DIMES and available for access, the system includes infrastructure to support requests by users. Requests include downloading or emailing archival record citation information, asking to view records onsite, or requesting digital copies. In an effort to simplify and improve on our previous approach to managing requests that had been embedded in the DIMES application, we created a separate application called the Request Broker (Arnold 2021) that serves as a layer between the DIMES frontend and our request management system, Aeon. Using API endpoints to integrate systems, the Request Broker can fetch and format data that researchers using DIMES do not need, like physical locations, directly from ArchivesSpace. It also implements a pre-request check to limit unfulfillable requests based on factors like access restrictions or if a digital version is already available with the goal of eliminating unnecessary labor by Reference staff in fulfilling requests. Like the modularity of the data pipeline infrastructure, this separation of functions improves error troubleshooting and flexibility in the relationships between systems.

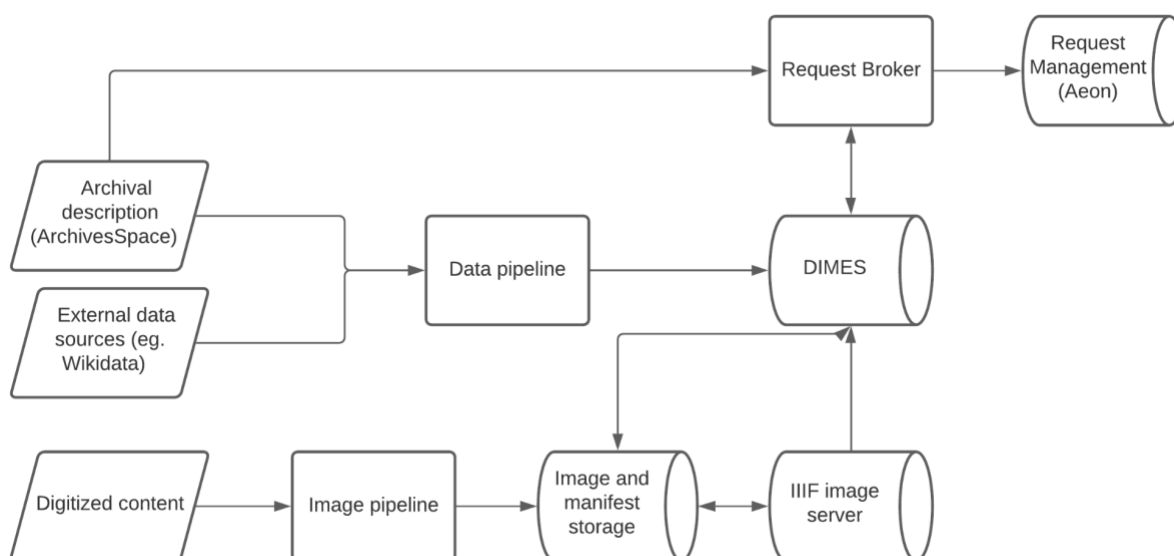


Figure 1: Diagram representing the basic DIMES infrastructure including the data pipeline, image pipeline, and Request Broker.

DACS Statement of Principles

Uncovering relationships

Our aims for uncovering relationships through a redesigned archival discovery and delivery system emerged from trends in the archival profession reflected in the Describing Archives: A Content Standard (DACS) Statement of Principles. The introduction of the revised DACS principles emphasize that archival description exists to “facilitate the use of archives by people,” and that “records, agents, activities, and the relationships between them are the four fundamental concepts that constitute archival description” (SAA TS-DACS 2021, Introduction). The principles articulate that the connections between these concepts can convey meaning and elucidate networks of interactions that would not have been apparent from the records’ content alone. Archival description exists to enable users to uncover relationships and so must the discovery systems that enable users to access that archival description.

When the RAC began defining requirements for a new archival discovery system, this need to leverage our description to surface networks between records, agents, and activities to enable user exploration was a primary consideration.

Our previous discovery system had been organized around searching within finding aids where the user could see an associated container inventory, but this traditional document-based finding aid approach offered limited means for uncovering relationships in the archives. While this would be true for any institutions' collections, it is especially relevant to philanthropic records, the RAC's primary holding. Insular yet global in scope, the world of philanthropy is both interconnected and far reaching, with a thread of ideas and people linking the enterprise together. The records are no different. Searching the records by creators alone either leads to an overwhelming amount of search results, with dozens of individual finding aids linked by name only, or, potentially, provides too little information to continue researching. This is compounded by the size and scope of these philanthropies, which often function on substantial endowments with a large number of staff members. Researching these records requires as many access points as possible.

In addition to the content of the records, a new discovery system was also a necessary adaptation to our improved processing workflows. In the past few years, the RAC has implemented a standardized archival processing workflow organized around processing by accession in which archivists provide DACS single-level minimum descriptions in order to make records publicly available more quickly and eliminate processing backlogs. Due to this processing strategy, and because we accession records on an ongoing basis from the same institutions, creators' records tend to be distributed across multiple related resource records in ArchivesSpace rather than in single collections, hence the utility of the arrangement maps described above. Adding a note describing related resources to a finding aid is insufficient for these circumstances because a similar note would need to be created for each resource record in this interconnected web of collections and every note would need to be updated when a new collection is added. The single finding aid note also requires researchers to start with a particular resource record first and then move out to related resources rather than provide them the option of starting with the network of related resources and the creator that ties them together.

In its fourth Principle of Archival Description, DACS states that "meaning in archival records is revealed through their contexts as much as through their contents. Archivists expose contextual significance by describing records, agents, activities, and the relationships between them" (SAA TS-DACS 2021, Principle 4). Although agents as records' creators were present in our previous discovery system, there was limited functionality to explore these agents as an access point for research beyond a creator. The agent relationship users could access was the one between a creator and a single collection within the finding aid for that collection, and the presentation of that relationship cast the agent as merely a detail in the finding aid, rather than as a person, family, or organization with its own descriptive data and its own network of relationships. In DIMES, the RAC sought to creatively leverage our existing archival data to provide new pathways to discovery.

Building relationships

DIMES was designed to enable the discovery of relationships in archival description, but also to build relationships between people. Archival description and the technical infrastructure and interfaces that provide access should be user-centered, and as the DACS Principles note, "users include not only those outside the repository, but the repository's own staff" (SAA TS-DACS 2021, Introduction). Therefore,

building something new was an opportunity both to enhance access for users and to collaborate with colleagues to design and improve processes in ways that value our labor and expertise.

Building on the DACS Principles, particularly the articulation in the fourth principle that “archival description that is rooted in ethics will produce a richer researcher experience,” the RAC has articulated six Guiding Principles for Archival Access (RAC 2019a), including that records and description are open by default with transparent restrictions, user access is self-directed to allow users to choose the level of mediation they require, user data collection and data retention should respect user privacy, access is generative and supports multiple pathways and modes of inquiry, our user interfaces are responsive and accessible, and that Reference staff have the infrastructure support to focus on their core activities. These guiding principles seek to root decisions about description, reference processes, and technical infrastructure in a framework that promotes responsibility, accountability, equity, and accessibility.

The collaborative work of building DIMES

The RAC contracted with a design agency, ondesign (<https://ond.com/>), to design the DIMES website based on initial RAC wireframes, information architecture maps, user research insights, and inspiration from the ArcLight Project (Stanford Libraries, n.d.) and many other existing discovery interfaces used by allied institutions and colleagues across our communities of practice. In collaboration with the other program areas, the RAC Digital Strategies Team led the project strategy and development work for the backend infrastructure and frontend website, with an emphasis on building and contributing to open-source systems that maximize interoperability and use community-maintained standards. All code and documentation are available in RAC GitHub repositories (RAC, n.d.). The project strategy included leveraging archivists' knowledge and skills to clean up our archival data, incorporating staff members' perspectives and expertise through participatory research and design, and communicating to staff members information which would resonate with their work while also providing forums for them to ask questions and issue feedback.

Agents data cleanup

In order to provide the functionality intended for DIMES and make full use of the Rockefeller Archive Center's data model, the RAC had to first prepare its data to fit this enhanced relational context. The RAC's Processing Team in collaboration with Digital Strategies completed several different projects targeting, at a macro level, key data elements as part of a larger data cleanup initiative (Young 2020). These elements included agent records (Berish, Martin, and Young 2020), dates (Berish 2020), and access restriction notes (Martin 2020). For the purposes of this case study, we will focus specifically on the agents cleanup work because of the key role agents play in the relational presentation of data within DIMES, but much of the rationale and collaborative processes driving the agents project was shared across the larger data cleanup effort.

The Rockefeller Archive Center's agents data was not prepared for the changes envisioned for DIMES partially because our archivists may not have fully considered the utility and meaning of agents as objects. As described earlier, agents in the old DIMES served as mere details within finding aids, and this mode of thinking likely became part of archivists' practices for linking agents to resource records. The finding aid was centered at the expense of the agents and their relationships. Understanding the significance of agent records in conveying contextual relationships to researchers helped clarify the

various issues impacting our agents data in ArchivesSpace. These issues included duplicate agent records meant to represent the same entity, inaccurate and incomplete data within agent records, and an overall inconsistent and non-standardized approach to creating agents. Furthermore, our repository held a massive amount of agent records linked to file-level objects in our Ford Foundation grants and catalogued reports collections that were exported from systems maintained by the Ford Foundation, one of our donor organizations. These agents had a disproportionate impact on our overall agents data because archivists at the RAC do not link agents to the file level as part of the regular processing workflow, and they would not have lent themselves to the type of relationships DIMES was intended to reveal.

As the first collaborative venture in the larger ArchivesSpace data cleanup initiative, the agents project began the relationship building that would empower processing archivists to work within and across team lines as well as to trust their own judgement in decision making. After exporting all of the RAC's agents data into CSV files for each agent type (person, corporation, and family) with help from the Digital Strategies Team, a group of processing archivists investigated the various problems impacting agent records in the hopes of devising an automated approach to comprehensively eliminate all of the duplicate records from our repository. These archivists had developed some competency with Python scripting through previous collaborations with Digital Strategies, and they knew that in order to leverage Python in this scenario, they needed to discover a pattern amongst the data that a script could understand in order to identify the correct records for deletion. Unfortunately, the issues affecting our agents data were too complex and various for the processing archivists to decipher a pattern around which to develop a script. In this moment, the archivists needed to have confidence in their own assessment of the situation and trust that the obstacles they confronted were not due to lack of technical expertise. By valuing the processing archivists' specific viewpoint rather than merely assigning them work, the DIMES project gave them the space to act upon their own judgement and opt for a manual approach instead. Using a workflow centered around the Enhanced Agent Merging Function in ArchivesSpace, the archivists were able to successfully merge or delete 6,704 agent records which was about 18% of all of the RAC's agents. The processing archivists would later have the opportunity to flex their Python skills when tackling the issue of the file-level agents in the Ford Foundation grants and catalogued reports collections. Working with Digital Strategies, they successfully wrote a script (RAC 2020) that unlinks all agent records from file-level archival objects within an indicated resource record. The script unlinked a total of 82,041 agents across 18 Ford grants and catalogued reports collections. These unlinked agent records were later deleted from the RAC's repository, completing the preparation of agents data for inclusion in DIMES.

Participatory research and design

A new discovery system and conceptual discovery model meant changes in the way staff at the RAC would do their jobs, particularly staff members focused on reference services, digitization work, and reading room retrieval who work most directly with researchers. The RAC Digital Strategies Team provided technical and strategic leadership for the project, but the goal was always to build *with and for* all teams across the organization, not to surprise people with an entirely new system that they would then be forced to use with no support. To value and incorporate the labor, expertise, and perspectives of staff users and build collaboratively, the project included contributors from across the organization to help define project requirements, conduct user experience research projects, join participatory design activities, and receive relevant training opportunities to support this work.

As part of the larger Project Electron initiative over the past four years culminating in DIMES, a large number of staff members from across the organization contributed to various work including user stories that helped define initial project requirements; participated in a card sorting activity to categorize user stories, define user groups, and create project personas (RAC 2019b); joined scenario mapping activities to improve understanding of existing archival processes and workflows; created service blueprints to articulate under-documented reference processes, surface staff labor that was often invisible to other staff, and identify pain points that DIMES might help address; participated in data modeling workshops to learn about and help draft the new data model for the project (Galligan 2018); created conceptual site maps for archival discovery to think through the DIMES website's information architecture and user flow; and contributed to ongoing usability testing studies of the DIMES website. Many of these methods come from the user experience design (UX) field and were new to the RAC but are part of a broader organizational strategy to develop UX expertise and approaches. The RAC does not have a UX team, but Digital Strategies, a team of four, defines improving user experience as one of its core activities. These cross-team collaborative UX projects and activities have multiple relationship-building benefits:

1. Gain knowledge from the activity and its artifacts to design and improve the user experience.
2. Encourage stakeholder ownership and investment in the success of the project through participation and contributions to the work.
3. Enable participants to contribute to project development without writing code.
4. Spread knowledge about UX methods and user-centered approaches across the organization.

Usability testing and the Observers Team

The DIMES usability testing program (Sistrunk 2021b) serves as a salient example for this case study in how UX methods can build relationships that enhance usability, transparency, and collaboration in line with the DACS Principles and RAC Guiding Principles for Archival Access. The RAC's approach to usability testing is based on Steve Krug's *Rocket Surgery Made Easy: The Do-It-Yourself Guide to Finding and Fixing Usability Problems* (2010) which emphasizes lightweight and iterative testing and scales well in our context of limited UX resourcing. We had prior expertise and formalized templates (RAC 2021i) for creating tasks, running tests with users, and debriefing with test observers to identify usability problems and propose fixes. The DIMES usability testing study was the RAC's first ongoing study that included testing at all project stages from design prototypes through development and post site launch.

The usability study planning and facilitation was undertaken by the Digital Strategies Team, who worked closely with an "Observers Team" consisting of one representative from each of the RAC program areas. The usability testing facilitator designed and facilitated four rounds of site testing during the development process and one after site launch, with each round consisting of a pilot test with an RAC staff member and three test sessions with an external user. The first round was conducted with researchers onsite in the reading room and focused on testing simple prototypes to compare two site concepts that were under consideration. Subsequent testing focused on specific site features, and sessions were conducted via moderated remote testing and recorded with the user's permission. After each round, the Observers Team watched the recordings and debriefed with the facilitator to identify usability issues and what might require further testing. The facilitator met with the DIMES developer,

who also observed the test sessions, to determine how to fix the issues. Finally, they reported the results to a group of staff from the Reference Team and shared a site demo. This group's deep collections domain knowledge and understanding of current request, retrieval, and digitization processes made their feedback and awareness of the results essential as the staff who work most closely with the researchers who use DIMES.

In evaluating this approach, there were two important challenges. First, the COVID-19 pandemic forced the RAC to shift from in-person testing to remote testing. This necessitated some technical and strategic adjustments but was ultimately an opportunity to expand the recruitment of participants beyond our reading room to include researchers located outside of the United States, researchers who did not speak English as their first language, were not familiar with our collections, had never conducted research in archives, and who came from non-academic contexts. A broad definition of who our users are and might be in recruiting participants in user experience research can support design decisions that privilege equitable access and accessibility. With this in mind, future usability testing should also include users with disabilities who use assistive technologies.

The second challenge to this approach is deciding how to fix observed problems within DIMES. In the RAC context, there is an overlap of expertise between developers and UX practitioners, and because the site developer worked closely with the usability project lead within the Digital Strategies Team to interpret the tests and implement site changes, communication and responsive action was not a challenge. However, translating observed usability challenges to design solutions and implementing them in code can be a barrier to this iterative approach when working with larger teams and/or in contexts with more distance between these roles.

Internal launch and rollout program

Two months before publicly launching DIMES, the RAC released the site internally for staff members to access along with a Google Form to provide structured feedback. The goals were to help identify bugs in the system, allow people to explore and test the site with their individual workflows, and give them the opportunity to ask questions and provide feedback before the launch required them to use the site in their day-to-day work. As this case study has detailed, there was wide organizational involvement in creating DIMES, so there were no major surprises. However, change inevitably introduces new challenges, and seeing demos and updates on project progress and even contributing to the work is distinct from understanding and being comfortable using the resulting system, particularly for those whose job functions are tied so closely to that system.

To kick off the internal launch, all RAC staff members were invited to attend a “rollout program” presentation and Q&A to provide an overview and refresher about why the RAC built DIMES, how all teams had contributed to its creation, details about its backend architecture, how archival data cleanup and enhancement work benefited discovery, and finally a demo highlighting important features and addressing changes that would impact existing processes and conceptual models. The team of presenters were recruited strategically as a group of archivists from different teams who could speak from a range of experiences from defining broad project vision and goals, architecting underlying infrastructure, conducting UX research, to improving archival data. The demo was conducted by a member of the usability testing Observers Team who is a reference archivist, which was a particularly successful approach to communicate important changes that related to this essential staff user group of DIMES.

The primary challenge of the rollout program was to communicate important concepts that were relevant to peoples' work without overwhelming colleagues with technical information that can act as a barrier to engagement. Additionally, while the feedback form received about 50 responses through the month of the public launch with useful insights and bug reports, the format meant that this feedback and the responses did not promote transparency around reported problems and responses. Using a Kanban board or similar tool to track and share feedback and/or changes more broadly within the organization may be a more effective future approach.

Maintenance and enhancement

DIMES was born in response to, and in anticipation of, the present and future needs of our user communities. DIMES is by no means a finished product, and we plan to enhance and refine DIMES in response to users' needs. Maintaining and continuing the work to build and uncover relationships in archival description and between people is essential to sustain the project and each other.

Relationships in archival description

Opportunities for collaborative projects to prime our data and the system's infrastructure for representing archival relationships continue to emerge. As previously described, the data pipeline architecture can draw on external data sources like Wikidata, and with the newly released expanded agents module in ArchivesSpace v3.0, the RAC can act on its plans to use these sources to enhance agent records. Processing archivists will build agent profiles for each agent type that will leverage the module's capacity for better expressing relationships by linking to external data sources as well as by defining the relations that connect some of our agent entities to one another such as the familial relationships between the Rockefeller family person agents. Processing archivists will also tackle the next data element for the ArchivesSpace data cleanup project: subjects. Borrowing the approach from the agents initiative, they will use the merging function in ArchivesSpace to bring order to the over 30,000 subject records in our repository. Our aims for this project are to make our subjects more useful for search within DIMES and to create more meaningful relationships between our subjects and other objects in the RAC data model like collections and agents.

The work planned for both the agents and subjects data in ArchivesSpace will intersect with the RAC's culturally competent description (CCD) initiative, a program developed to make our description more inclusive and highlight peoples and histories that have been underrepresented because of the role agents and subjects play in describing people and their records. The new DIMES has given increased significance to both data objects, and within its relational presentation of archival data, agents and subjects can be employed to better articulate the power relationships at work within records of philanthropic organizations, one of the objectives of our CCD initiative. We will also bring an inclusive and reparative description framework to our agents and subjects work in order to resist the valorization of philanthropists and foundations, bring attention to the contributions of people from marginalized backgrounds, and better represent grantees and communities served.

Relationships between people

Aside from revealing relationships amongst our archival collections and data, the RAC's ongoing development and maintenance of DIMES post-release has the potential to further cultivate and nurture the relationships amongst RAC staff. Sustainable infrastructure is one of the stated aims of Project

Electron, and for infrastructure to be sustainable there must be people capable of and invested in maintaining the system. By incorporating the perspectives and expertise of staff from across our different program areas, the DIMES project prepared our staff members to contribute their specific knowledge to the maintenance of DIMES through technical repair, error reporting, and articulating areas for improvement that impact their work. An example of this cross-team collaborative approach to maintenance is the data pipeline troubleshooting team which is tasked with diagnosing and resolving issues impacting data movement from ArchivesSpace to DIMES. The team consists of representatives from Digital Strategies, Processing, and Information Technology, and different members of the team are assigned tasks based on their particular strengths.

As the primary staff users of DIMES, RAC reference archivists are an essential nexus. The Reference Team's experience of and contribution to DIMES will rely on relationships, both with the researchers and with each other. Relationships with the former will provide a window into the researcher experience, while our relationships amongst ourselves will help us improve our services, workflows, and DIMES itself.

This is all the more relevant as the RAC staff faces the reference challenges brought upon by the COVID-19 crisis. At the time of writing, the Reference Team does not have the usual access to the researchers and their processes because the RAC is closed to in-person research. Whereas the Reference Team's pre-COVID procedures relied heavily on one-on-one in-person interviews, the current situation denies us the opportunity to visually walk through the researchers' encounter with DIMES. Additionally, the team's previous reference interactions also relied heavily on assisting the researcher in understanding the cross-collection nature of philanthropic records, and the extent this message is translated to users of DIMES is uncertain. Are the researchers approaching DIMES with the expectation that the search results will be organized as if run through Google, with a list of item-level search results? How are they encountering and creating the context of the records?

The Reference Team has also encountered an unexpected tension between the goals of DIMES and the researchers' immediate experience. In an attempt to provide equitable and flexible access to researchers, researchers are encouraged to explore materials without fear of copyright infringement. However, reproducing these materials may not come so easily, as publishers are less willing to assume risk. As researchers continue to find more material from DIMES, it is possible the RAC will receive more inquiries regarding copyright and permissions.

For all these uncertainties, there are also opportunities for learning, creating, and expanding best practices in our reference interactions with researchers. The RAC staff is currently exploring the potential of a "reference knowledge share", where archivists who do reference work can talk through what we have observed about DIMES thus far, and how they have answered questions from researchers about the system. Though reference archivists are no strangers to discussing individual requests among themselves, this will be the first formal meeting of this kind and will involve a radical vulnerability as they reveal the inner workings behind their individual approaches.

In addition to this narrative-based strategy session, the Reference Team plans to tap into the expertise of a Team member who, thanks to another intra-departmental initiative (Sistrunk 2021a), is knowledgeable about existing RAC web analytics tools and DIMES analytics data in particular. From this, the Reference Team hopes to gain insight into the experiences and behavior of those who may not

choose to reach out to us directly. Before the pandemic, the Reference Team processes centered on the in-person researcher visit or individually-submitted duplication requests. Listening and responding to data gleaned from groups of users, and not just individual researchers as experienced through the affective interaction with a reference archivist, will be a new experience, but it will be an opportunity to build on what we learned, and the skills we gained, from usability testing. It might also prove to be more important as the RAC move towards a post-COVID world — it's entirely possible that users will rely on remote services, either by necessity or by way of their new modes of working.

The Reference Team will also need to adapt to new modes of working. Pre-pandemic, the day-to-day operations and busy reading room did not always leave time for reflection. As the RAC reopens the reading room and responds to an influx of digitization requests, it will be necessary to build in the time to learn from each other. This restructured concept of time will apply not just to the Reference Team, but across the RAC, as we, like the researchers, think through and adapt to a changed world.

Conclusion

Relationships, both amongst the staff and across our collections, played a key role in the creation, implementation, and maintenance of DIMES. What started as a project with common professional values and principles developed into a set of shared systems, workflows, and skill sets that encourage transparency, engagement with technology, and cultural and organizational change and growth. From this work, we have grown to understand the importance of the continued maintenance of the relationships that built and will sustain this system. We strive for our work in both building and enhancing DIMES to serve as a model for other institutions interested in developing their own systems for archival discovery and delivery.

Acknowledgements

The authors have represented this case study from our own perspectives, but DIMES, as part of the broader Project Electron initiative, was a truly collaborative effort of the entire Rockefeller Archive Center staff with leadership from President Jack Meyers and Director of Archives Robert Clark. We acknowledge our colleagues who have contributed their labor, knowledge, and innovative ideas over many years. The RAC contributors whose work was specifically described in this case study include the Digital Strategies Program members Bonnie Gordon, Patrick Galligan, and Assistant Director Hillel Arnold; Processing Program archivists Amy Berish, Katie Martin, and Assistant Director Robert Battaly; the usability testing Observers Team including Amy Fitch, Reuben Slater, and Rachel Wimpee; and the usability testing Reference Program consultants Kanisha Greaves, Bethany Antos, and Assistant Director Michele Beckerman.

References

- Arnold, Hillel. 2021. "Request Broker: An Overview." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated June 3, 2021. <https://blog.rockarch.org/request-broker-overview>.
- Berish, Amy. 2020. "ArchivesSpace Clean Up: Adding Structured Dates to an Entire Repository." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated April 28, 2020. <https://blog.rockarch.org/archivesspace-cleanup-dates>.
- Berish, Amy, Katie Martin, and Darren Young. 2020. "ArchivesSpace Clean Up: Wrangling Agents in the Wild." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated May 5, 2020. <https://blog.rockarch.org/archivesspace-cleanup-agents>.
- Galligan, Patrick. 2018. "Modeling for Project Electron." GitHub repository. Last updated April 6, 2018. <https://blog.rockarch.org/modeling-for-project-electron>.
- Krug, Steve. 2010. *Rocket Surgery Made Easy: The Do-It-Yourself Guide to Finding and Fixing Usability Problems*. Berkeley, California : New Riders.
- Martin, Katie. 2020. "ArchivesSpace Clean Up: Removing Legacy Access Restriction Notes." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated April 30, 2020. <https://blog.rockarch.org/archivesspace-cleanup-notes>.
- Rockefeller Archive Center. n.d. Rockefeller Archive Center Organizational GitHub Repositories. Accessed July 1, 2021. <https://github.com/RockefellerArchiveCenter>.
- Rockefeller Archive Center 2019a. "Guiding Principles for Archival Access." Project Electron. Last updated October 16, 2019. <https://projectelectron.rockarch.org/archival-access-values>.
- Rockefeller Archive Center 2019b. "Rockefeller Archive Center Personas." GitHub repository. Last updated May 1, 2019. <https://github.com/RockefellerArchiveCenter/personas>.
- Rockefeller Archive Center. 2020. "remove_agents.py". scripts GitHub repository. Last updated January 29, 2020. <https://github.com/RockefellerArchiveCenter/scripts/blob/base/archivessnake/remove-agents.py>.
- Rockefeller Archive Center. 2021a. "argo." GitHub repository. Last updated June 29, 2021. <https://github.com/RockefellerArchiveCenter/argo>.
- Rockefeller Archive Center. 2021b. "DIMES: The Online Collection and Catalog of Rockefeller Archive Center". Last updated April 3, 2021. <https://dimes.rockarch.org>.
- Rockefeller Archive Center. 2021c. "IIIF Pipeline." GitHub repository. Last updated June 12, 2021. <https://github.com/RockefellerArchiveCenter/iiif-pipeline>.
- Rockefeller Archive Center. 2021d. "pisces" GitHub repository. Last updated May 17, 2021. <https://github.com/RockefellerArchiveCenter/pisces>.

- Rockefeller Archive Center. 2021e. "Project Electron: Building Infrastructure for Data in Motion." Last updated June 30, 2021. <https://projectelectron.rockarch.org>.
- Rockefeller Archive Center. 2021f. "RFC 005 - Arrangement Maps." project_electron GitHub repository. Last updated June 30, 2021. https://github.com/RockefellerArchiveCenter/project_electron/blob/base/rfcs/005-arrangement-maps.md.
- Rockefeller Archive Center. 2021g. "Simplified Data Model." rac_schemas GitHub repository. Last updated March 2, 2021. https://github.com/RockefellerArchiveCenter/rac_schemas/blob/base/Simplified_Data_Model.png.
- Rockefeller Archive Center. 2021h. "scorpio." GitHub repository. Last updated June 29, 2021. <https://github.com/RockefellerArchiveCenter/scorpio>.
- Rockefeller Archive Center. 2021i. "Usability Testing at the Rockefeller Archive Center." Rockefeller Archive Center Documentation. Last updated May 18, 2021. <https://docs.rockarch.org/ux-templates>.
- Society of American Archivists Technical Subcommittee on Describing Archives: A Content Standard (SAA TS-DACS). 2021. "Statement of Principles," *Describing Archives: A Content Standard*. Version 2021.0.0.2. https://saa-ts-dacs.github.io/dacs/04_statement_of_principles.html.
- Sistrunk, Hannah. 2021a. "Breaking Isolation Through Cross-Program Collaboration." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated May 17, 2021. <https://blog.rockarch.org/analytics-collaboration>.
- Sistrunk, Hannah. 2021b. "Leveraging Collaborative UX Research, Design, and Testing Strategies to Operationalize UX Practice." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated March 1, 2021. <https://blog.rockarch.org/dimes-ux>.
- Stanford Libraries. n.d. "ArcLight." Stanford Libraries Projects & Innovations. Accessed June 30, 2021. <https://library.stanford.edu/projects/arclight>.
- Young, Darren. 2020. "ArchivesSpace Clean Up: An Outline." *Bits & Bytes: The Rockefeller Archive Center Blog* (blog). Last updated April 23, 2020. <https://blog.rockarch.org/archivesspace-clean-up-an-outline>.

Access is People: How Investing in Digital Collections Labor Improves Archival Discovery & Delivery

Stephanie Becker, Anne Kumer, and Naomi Langer

Abstract: Archivists are increasingly expected to provide remote digital access to their physical collections in order to meet contemporary research needs. The labor involved in creating and stewarding digital collections, however, is often seen as a support role to the stewardship of physical collections, which causes inconsistent and unsustainable digitization projects and contributes to hierarchical communication structures and archival labor precarity. In this paper, the authors consider various stakeholders in creating digital collections - researchers, library administrators, archivists, and digital collections staff - and argue for a shared stewardship approach to digital collections project management and policy development through the case study of their own experiences forming and serving on a Digitization Governance Committee at the Kelvin Smith Library of Case Western Reserve University.

Introduction

In cultural heritage institutions, digital collections labor such as digitization, metadata, and repository work are often carried out by archivists responsible for stewarding physical collections, term-based employees filling archival labor gaps, or student and volunteer positions. In this paper, we will examine how a shared stewardship model for archival collections, where all labor is valued as a core function, can further access to unique collections and foster a more equitable work environment for those who make access possible. The lack of financial investment in full-time digital collections staff signals that this labor is understood to be a secondary service in the stewardship of physical collections. Even when dedicated digital collections staff is present, it is common for those individuals to work in isolation from archivists who carry out physical collections labor such as acquisitions, processing, and reference work. Our goal is to think openly about problem-solving and expand beyond established methods and approaches to work that continually enforce labor precarity, hierarchical structures, and inequitable resource allocation.

Throughout this paper, we will assess stakeholder needs and highlight our own experiences of working together, alongside other colleagues, at Case Western Reserve University's (CWRU) Kelvin Smith Library (KSL), an academic research institution in the United States. Our positions within the library fall outside the Scholarly Resources and Special Collections (SRSC) team which is composed of archivists who steward the library's physical collections. Naomi Langer and Stephanie Becker are part of the Freedman Center for Digital Scholarship team and are responsible for digitization and repository work. Anne Kumer is responsible for digital collections metadata and is part of the Acquisitions & Metadata team.

These distinctions are noteworthy because the way that labor is organized within any given institution will influence the impact of that labor and the personal success of those who provide it.

In a recently published OCLC research report, *The Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections*, the authors note that “in many institutions, those tasked with building collections, are separate from those tasked with the ongoing stewardship work of collections,” (Weber et al. 2021) and that this in turn has negative consequences on archivists who are facing large backlogs of unprocessed materials. They lay out a total cost of stewardship framework for thinking about acquisitions and collections care holistically, so that decisions can be made from a more equitable standpoint. We argue that this idea can be expanded to digital collections staff who are often separated from the work of archivists. Decisions made about processing and description impact the work of digital collections staff - especially as it pertains to the differences in best practices, interoperability of systems, and the use of available resources to provide access to both the physical and digital collections in our care. By including all perspectives, our institutions will meet a wider variety of stakeholder needs and in turn foster an equitable decision-making process that results in sustainable collection policies.

Identifying Stakeholders

Stakeholder needs largely influence what administrators at cultural heritage institutions will spend money on regarding the stewardship of their special collections. For unique digital collections at KSL, we have identified researchers, library administrators, archivists, and digital collections staff as our main stakeholders. These groups work in tandem and therefore clear communication is needed to set expectations and respond to changing needs over time. During the past year, we have seen a shift from in-person to remote learning during the height of the COVID-19 pandemic, and as a result, the needs of our stakeholders have changed. Given the shift toward a fully online research lifecycle, we have identified that now is an opportune moment to re-examine stakeholder needs and think creatively about the labor required to meet them.

Researchers

After a full year of remote engagement, researchers expect consistent and easy access to resources that do not require learning a multitude of different processes and policies just to access content. At KSL, we have an institutional repository, library catalog, patron request system, and online finding aids.¹ While each of these systems provides a necessary internal function, externally, it can be confusing to manage a personal user account across systems and know where to begin your research. Through our patron interactions, we have also learned that researchers not only want remote access to materials, but they also want the ability to download those materials for a wide variety of purposes. Even if a researcher is able to visit the special collections reading room to view objects of interest, they leave wanting to take copies of those objects with them. As research methods evolve, researchers need access to high quality digital objects and descriptive metadata that they can leverage with text mining, mapping, data visualization, and other digital scholarship tools. In stewarding physical collections, archivists are laying the foundation for digital collections staff to meet the needs of researchers who want to access and leverage collection objects in a digital environment. Furthermore, creating and

¹ This is not a full list of systems in use at KSL, but rather a list of systems related to digital collections work. These include Islandora 7, Sierra, Aeon, and ArchivesSpace.

sharing high quality digital objects and descriptive metadata requires an entirely different set of expertise and resources than what is typically available in archival units.

Library Administrators

Since we work in an academic institution, our perception of library administrators' main interests is to meet the needs of donors and researchers, with a heavy emphasis on faculty. They also require a way to sustainably maintain ongoing strategic and operational initiatives from a financial and policy perspective. At KSL, our yearly budget is determined by CWRU administration and supplemented with one-time donations and ongoing endowment funds. Financial spending decisions are influenced by the needs of donors and researchers who advocate for the forward-facing resources they require, and not the long-term infrastructure and staffing needed to provide those resources. For example, potential collection donors can negotiate digitization and online access as a stipulation for acquiring their materials. If the availability of institutional resources does not cover sustainable funding of digital collections work, then archivists are left with the burden of meeting the donors' needs without additional support from library administration. While grant funding can be used on short-term labor for a specific project, such as digitizing a recent acquisition, the dependence of temporary labor "negatively affects everyone involved — the archivists, institutions, collections, donors, and users" (Dean et al. 2018). In 2018, a group of temporarily employed archivists at the University of California Los Angeles (UCLA), published an open letter to their library administrators outlining the negative consequences of hiring archivists on temporary contracts to perform ongoing work. The authors cite consequences such as low staff morale, lack of valuable institutional memory, and the diverting of limited resources to recruitment and training. Any attempts by administrators to avoid such negative consequences, then must include financial transparency with their stakeholders about the resources on-going digital stewardship work requires.

Archivists

Amidst pressure to respond to both changing research and library administrator needs, without full-time digital collections staff, archivists are expected to possess the time and expertise to execute complex digitization workflows that include applying consistent description, preservation of digital surrogates, and maintaining a repository or alternative online access point. At the same time, many institutions have decreased full-time staffing resources in archives and special collections departments. The precarity of archival labor, which includes those who specialize in digital collections, has previously been documented as a field-wide problem with numerous consequences to those who carry out the work. In a 2019 study on the experiences of precarious employment in Canadian libraries, respondents expressed negative personal effects including "financial and psychological vulnerability, difficulties with physical and mental health, difficulty pursuing social activities, and choosing to delay significant life decisions," (Henninger et al. 2019), all of which are disproportionately faced by marginalized groups such as women and people of color. As budgets shrink and archivists are increasingly responsible for more, the only support they can secure is transient student labor, grant-funded processing staff, and volunteers, all of whom are temporary support for ongoing operational work. Where a department may have once had two full time archivists who divided work based on function — reference, processing — it has grown increasingly common to see one archivist tasked with reference, processing, metadata remediation, and donor relations. Archivists need colleagues who can focus on the digitization and online access of collections so that they can focus on working with patrons and processing new acquisitions.

Digital Collections Staff

The issue of labor precarity has a long-term impact on how digital collections are created and maintained over time. Researchers and library administrators can desire quick remote access to archival materials, but the labor behind making that happen involves more time and cross-departmental effort than often assumed, at the expense of both the people performing that work and the quality of the eventual digital product. The need for job security means that those hired into temporary positions are subject to leave at any moment. When one project is passed through many different hands, it becomes impossible to apply standards and achieve any form of consistency, and this in turn has negative consequences for researchers trying to access and use digital collections. When hired as term-based employees to carry out digitization, metadata, or repository work, the people in those positions are usually supervised by a full-time archivist, which enforces the narrative that digital collections staff are hired to support the core work of collections. Supporting roles then, are left out of archival decision making and policy changes, which creates a culture where staff feel comfortable with vertical but not lateral communication on work that demands collaboration. Perhaps the most overlooked consequence of labor precarity is the reality that it costs cultural heritage institutions more money in the long run. Those lucky enough to secure full-time permanent employment in digital collections can attest that the first several years of any position is cleaning up old backlogs and previously digitized collections that were ingested into digital repositories and storage environments where they may be lacking context or preservation plans. Meanwhile, new work piles up and becomes yet another backlog, making it difficult for digital collections staff to truly succeed. Institutions that hire temporary staff to address backlogs, then pay numerous people for a short period of time to do the same work over and over again. What can be achieved with permanent staff is a set of sustainable policies that allow for consistent work to happen no matter who holds those positions. In other words, you can spend money for years on band-aids, or you can invest money in people who can create sustainable solutions that include ongoing maintenance plans and eliminate the need for re-doing work time and time again.

What Valuing Labor Looks Like in the Workplace

In May 2019, KSL's Digital Collections Manager established an internal Digitization Governance Committee (DGC) to provide oversight of digitization and digital collections project planning, policies, and workflows. As discussed in *A Vision for Kelvin Smith Library's Digitization Program* (Becker. 2019), which lays out the need for sustainable policies based on human action, our goal when digitizing collections is to "do it once, do it right". This keeps our fragile and rare objects from being handled more than necessary and saves library administrators money in staff time and energy that would otherwise be spent redoing past digitization and digital collections work. To achieve this goal, DGC needed member participation across several departments in the library to ensure that expertise in digital collections, preservation, metadata, physical collections, cultural heritage imaging, and library administration all have a voice. DGC members collaborated on writing cross-departmental workflows that we rely on for digitizing and providing access to collections. Furthermore, the Digital Collections Manager maintains an internal Google Site, accessible to all KSL staff, that serves as a central access point for posting meeting minutes and finalized DGC policies. The site allows for full transparency into how and why the committee made each decision in our resulting policies that we review and update on an annual basis. Archival objects either from our set Yearly Digitization Plan (also worked out by DGC) or requested by library patrons are pulled by an Archivist, retrieved by the Preservation Officer for a condition and handling review, who then brings the objects to our Digitization Lab for the Digitization

Technician to photograph and output the digital files. Our Metadata Librarian prepares item-level descriptive metadata for the Digital Collections Manager to pair with the digital files and ingests the objects into our institutional repository. The Preservation Officer re-inspects the physical object and then brings it back to its proper place in storage.

Applying Best Practices

Digital and physical collections work adheres to different standards and best practices that can cause pain points when digitizing and sharing collections online. Because employees with varying skills and knowledge sets are often spread across multiple teams in an institution, each of which has a different understanding of their field's best practices, and often a different level of authority over collection development, developing successful digital collections can be challenging. Mitigating pain points through collaborative policymaking enables collections staff to think through problems together and create long-term solutions instead of responding to problems as they arise. Together the DGC figured out how to make this process efficient for each party, and how best practices are translated throughout the overall workflow. Issues that may have arisen further down the line are better anticipated during planning, and cross-disciplinary issues that arise during project completion can be easily resolved. One instance of this is the differences in descriptive best practices between archival objects (typically at folder-level) and institutional repository objects (item-level) which was presented at a DGC meeting during early stages of workflow development. Through our conversations, we concluded that it was infeasible for archivists to provide item-level description and for the Digital Collections Manager to add objects to our repository without descriptive information. This best practice discrepancy led to several conversations where committee members worked through different metadata scenarios and created solutions for each that would be applied later on during the digitization process. Front loading the intellectual labor of planning will save time later during the ingest process and lead to fewer instances of rushed problem solving as unforeseen descriptive anomalies present themselves during collection processing, digitizing, and ingesting. It also means that staff turnover won't halt our workflows because a new employee can be on-boarded by DGC members with our sustainable set of governing policies.

Best practices for digital imaging also benefit from committee discussion, instead of being performed ad-hoc by individual library departments looking to make digital surrogates of their physical materials. Contemporary cultural heritage digitization has expanded beyond flatbed scanning and now requires an in-depth knowledge of advanced photography, imaging science, and industry-wide quality standards, like those set by the Federal Agencies Digital Guidelines Initiative (Still Image Working Group 2016). The policies created by the DGC, ensure consistency in collections imaging across departments and projects. Projects in our digitization lab come from Special Collections, University Archives, and local partner institutions. We photograph all collections objects at the same measured quality, using imaging targets and analysis software to check how faithful our digital surrogates are to their physical counterparts, as well as check the images' quality over time during ongoing digital preservation work. Using these standards allows us to meet our stakeholder needs by providing researchers with consistent high-quality images. It also expands the potential for archival discovery of our collections, by meeting quality standards for inclusion in national and international consortiums such as the Digital Public Library of America (DPLA) and the HathiTrust Digital Library. Our full-time Digitization Technician has the expertise required to meet these highly technical standards, and as a permanent staff member, can engage in professional development related to cultural heritage imaging, ensuring our work continues to meet stakeholder needs and is in line with that of our peer institutions.

Interoperability

The compatibility of computer systems, software, and programs to accurately work in tandem is an ongoing challenge for most libraries. A systems administrator who thinks holistically about how systems/people interact can mitigate many hindrances to interoperability, but we argue that interoperability is also reliant on transparent communication, mutual respect, an understanding of the larger workflow, and effective collaboration. In 2011, the EU funded project DL.org (Digital Library Interoperability, Best Practices and Modelling Foundations) convened a working group to identify and investigate interoperability challenges as they relate to digital libraries and collections work. The working group expanded beyond systems interoperability to include library policy on an organizational and semantic level: “This kind of interoperability takes place at a high (organizational) level, and it is then instantiated at a process level - whether those processes are being handled by human or machine. In terms of standards, policy interoperability is a step beyond policy standardization and is crucial to achieve useful interoperability between real-world digital libraries” (Innocenti et al. 2011). Though this example addresses interoperability among different institutions and libraries, the same methods can be applied to achieving interoperability among departments within one institution. Ensuring interoperability requires knowledge sharing and comprehension of workflows that are adjacent to any one person’s duties in collection stewardship. All of these provide a collaborative network where, ideally, all parties are consulted about benefits and pitfalls of various computer software and systems before those systems are purchased and implemented, as well as after they have been put into use. There is no one system that works best for all collection management and stewardship functions, but group problem-solving can go a long way in ensuring that each person / department’s needs are met.

Systems are also not effective without someone to manage and maintain both the system itself and the local content present within the system. Most institutions have a combination of vendor supported and open source systems. While vendored solutions may outsource some of the ongoing labor associated with these systems, library administrators must still identify a staff representative who is tasked with internal problem solving and maintaining the vendor relationship. A staff representative is also vital to internally maintain open source systems that require technical updates and engagement with the system’s open source community. Naturally, this person is well versed in their department’s or individual use and function of the system, but may not be entirely aware of malfunctions or deficits that affect adjacent workflows and overall collection stewardship. Boutique software and library systems are designed to fulfill a functional need, and not necessarily to work seamlessly with other software and systems (Foulonneau et al. 2008). Implementation of these systems happens almost solely within library administration and tech departments, while integration is left to collections stewards to troubleshoot as problems arise. During our conversations about discrepancies in descriptive best practices, DGC members identified that if the committee had been more involved in the selection and implementation of ArchivesSpace, we could have connected our finding aids to objects in the repository, perhaps eliminating the need to translate metadata from one standard to the other. Integration would also have impacted our metadata preparation and ingest workflows, by harvesting the work already done by our archivist, instead of our metadata librarian having to extract, transform, and ingest metadata into our repository system. A lack of advance consideration of these differing practices requires quick fixes that in the long run cost the library more money and staff time. So, even though the space held by monthly DGC meetings formalizes and legitimizes all functions of digitization through ongoing collaborative policymaking and group problem-solving; it doesn’t fully close the gaps caused by departmental disconnect.

The Digital Collections Manager initially convened the DGC as a space for drafting and maintaining intradepartmental workflows. With our core policies now in place, DGC members will now expand the committee scope beyond policy governance. We have recently had one of KSL's digital scholarship librarians, who also leads the library's instruction program, join our committee. With their expertise added to the established digitization workflows, the committee can begin brainstorming how we might integrate digitized collections into the classroom. The addition of other KSL staff not directly involved in the digitization workflow enables us to think about our work in different ways. For example, how might instruction needs impact the way we photograph and describe objects? Can we tweak aspects of our workflows to better serve librarians doing reference or interlibrary loan work? We plan to continue efforts around interoperability by engaging our colleagues, thinking about systems holistically, integrating currently disparate systems, and continuing to collaborate and address new and emerging issues as they arise.

Resources

Collaborations in libraries are often hindered by autonomous management of departments, inequitable allocation of resources, and a hierarchical staffing structure. All lead to disproportionate notions of the value of positions and work (chief among them, the damaging distinction between core and support roles), feelings of resentment, defensive communication practices, and low morale. Findings from a recent study examining dissatisfaction of digital stewards, introduced in a 2012 National Digital Stewardship Alliance (NDSA) survey, and again in a 2017 follow-up survey indicate lack of long-term planning and allocation of resources, lack of policy and decision-making authority, and lack of long-term commitment from leadership to be among the primary causes of stalled progress in digital initiatives and low staff morale (Blumenthal et al. 2020). While the formation of a larger, intradepartmental committee, such as DGC, addresses some of these concerns, all efforts require administrative support and advocacy to succeed.

The OCLC report mentioned earlier addressed the need to evaluate the total cost of collection acquisitions and management: "While we are accustomed to thinking of an annual collecting budget as a constraint on collecting, we are not as accustomed to thinking about our capacity to steward as a constraint" (Weber et al. 2021). This requires adopting a management model that moves away from an over reliance on term labor to perform a wide range of duties and towards one that invests in long-term professional development and sustainable full-time positions with a wide range of responsibilities. In 2016, four term employees, each from a different institution, presented the discrepancies between the projects they were hired to do, and the variety of tasks they ended up performing (Davis et al. 2016). Each presenter was hired to process or catalog a specific collection at their respective institutions, and all of them ended up filling larger digital collections labor gaps including digitization, metadata application, and repository management. Within this model that relies on soft money and temporary labor for operational work, progress is measured by short project timelines, grant deadlines, and fiscal year endings (Blumenthal et al. 2020). In other words, progress is measured by the presence (or absence) of resources over a short period of time, rather than the work completed over a long period of time. The additional tasks as assigned and outlined by each of the presenters are the catchall for the human element of staffing, too often under considered in planning for term projects: life emergencies that need tending to, staff turnover in favor of full-time positions elsewhere, and burnout from having to meet short deadlines with minimal resources. Acquiring resume building skills adjacent to the job one was hired for can be a good thing, but only when it is supported with adequate resources and guidance from colleagues.

In addition to offsetting operational costs, serendipitous donations, short-term grants, and other sources of soft money can free up the budget for professional development, but often only for those in the awarded department, allowing them to skill up at a faster rate than their colleagues. Those whose work is perceived as support, such as staff responsible for metadata creation and digital imaging, as well as staff who work to enable additional access, exposure, and care for collections are left out of consideration when these resources are obtained and distributed. Of the most sought after professional development resources — which allow workers to network, have access to continuing education, and enjoy opportunities for knowledge-share outside of their immediate institutional circles — are institutional memberships to professional associations. The memberships library administrators choose to pay for will be prioritized by the impact factor on both the institution and the number of staff it can serve. The lack of investment in full-time digital collections staff also means a lack of investment in associations that would benefit such staff. For example, administrators may fund institutional memberships to the Society of American Archivists (SAA) and the American Library Association (ALA), but not to the Digital Library Federation (DLF). This oversight prioritizes public facing positions perceived as core library functions, while denying others equal access to professional associations relevant to their work functions and their reduced status as support staff. It's tempting to codify resources as primarily financial, but positive staff morale — achieved by sustainable long-term project planning, equitable hiring practices, competitive compensation packages, and increased recognition of all collection stewardship work — is the most important resource a library can cultivate for long-term success.

Conclusion

Valuing all labor as a core function can help to establish a shared stewardship model for archival collections. We recommend slowing down in order to think through and create long-term sustainable policies that allow for increased discovery and access of archival materials while simultaneously creating a more equitable work environment for those who make access possible. There is a precedent in the delivery of archival resources of rushing to apply band-aid solutions in order to meet stakeholder needs. By hiding the realities of our labor, we are setting unrealistic expectations for researchers and library administrators. Being transparent about the work that goes into a research request is a positive action that will set a realistic standard for stakeholders regarding the time, expertise, and other resources it requires to meet their needs.

In the best case scenarios, valuing digital collections labor means hiring full-time digital collections staff with access to professional development and career growth opportunities. We recognize however, that many institutions do not have the necessary support to create full-time positions. In this scenario, we encourage library administrators to think creatively about how the available labor in your given institution is organized and valued. Increased open communication helps to prioritize the people doing the work and in turn the work being done, further empowering staff to apply group problem-solving skills to stewardship, no matter how small that group may be. Involvement of staff at all levels partially dismantles the hierarchy of top-down decision-making prevalent in cultural heritage institutions, but it does not negate the need for administrative buy-in and support for long-term stewardship. The Digitization Governance Committee purposely includes managers and administrators who don't directly participate in the digital collections workflow, so that they in turn can advocate for sustainable funding and staffing for all collections work.

The committee's success within our library is a step towards defining all labor as core labor, and redefining what project success looks like. It isn't so much that the completion of a single digitized collection is the metric for success, but rather the ability to apply that project's workflow to future projects. Rather than creating policies and workflows for one digitization project, the work we continue to do as a group holds greater value as a foundation for a digitization program, to be applied to many collections. Our sustainable policies save library administrators from spending their limited resources on short-term solutions and redundant labor, while honoring the skilled labor needed to make archival discovery and delivery possible.

References

- Becker, Stephanie. 2019. "A Vision For Kelvin Smith Library's Digitization Program". Edited by Jen Green and Roger Zender. Freedman Center for Digital Scholarship, Kelvin Smith Library, Case Western Reserve University (July). <http://hdl.handle.net/2186/ksl:2006056360>.
- Boyles, Christina, Anne Cong-Huyen, Carrie Johnston, Jim McGrath, and Amanda Phillips. 2018. "Precarious Labor and the Digital Humanities." *American Quarterly* 70 (3): 693-700. <http://doi.org/10.1353/aq.2018.0054>.
- Blumenthal, Karl, Peggy Griesinger, Julia Y. Kim, Shira Peltzman, and Vicky Steeves. 2020. "What's Wrong with Digital Stewardship: Evaluating the Organization of Digital Preservation Programs from Practitioners' Perspectives." *Journal of Contemporary Archival Studies* 7, Article 13. <https://elischolar.library.yale.edu/jcas/vol7/iss1/13>.
- Davis, Rosemary K.J., Rachel Mattson, Elliot McNally, and Alison Reynolds. 2016. "Grant Cycles, Deadlines, and Labor Advocacy: The Changing Work of Project Archivists" (Presentation, MARAC conference, Pittsburgh, PA, April 14–16). <https://doi.org/10.13016/M2X48F>.
- Dean, Courtney, Lori Dedeyan, Angel M. Diaz, Melissa Haley, Margaret Hughes, and Lauren McDaniel. 2018. "UCLA Temporary Librarians Open Letter." <https://doi.org/10.5281/zenodo.5159944>.
- Dohe, Kate. 2019. "Care, Code, and Digital Libraries: Embracing Critical Practice in Digital Library Communities." *In The Library With The Lead Pipe*. February 20. <https://www.inthelibrarywiththeleadpipe.org/2019/digital-libraries-critical-practice-in-communities/>
- Duplise, Genna. 2019. "Archivists on the Issues: Rare & Ephemeral: A Snapshot of Full-Time New England Archives Jobs, 2018-2019." *Issues & Advocacy* (blog), *Society of American Archivists*. May 28, 2019. <https://issuesandadvocacy.wordpress.com/2019/05/28/archivists-on-the-issues-rare-ephemeral-a-snapshot-of-full-time-new-england-archives-jobs-2018-2019/>.
- Federal Agencies Digital Guidelines Initiatives, Still Image Working Group. 2016. "Technical Guidelines for Digitizing Cultural Heritage Materials."
- Foulonneau, Muriel and Jenn Riley. 2008. *Metadata for Digital Resources: Implementation, Systems Design and Interoperability*. Chandos Publishing (Oxford).

- Hathcock, April. 2019. "Why Don't You Want to Keep Us?." *At the Intersection* (blog). January 18, 2019. <https://aprilhathcock.wordpress.com/2019/01/18/why-dont-you-want-to-keep-us/>.
- Henninger, Ean, Adena Brons, Chloe Riley, and Crystal Yin. 2019. "Perceptions and Experiences of Precarious Employment in Canadian Libraries: An Exploratory Study," *Partnership: The Canadian Journal of Library and Information Practice and Research* 14, no. 2. (September 23, 2019). <https://doi.org/10.21083/partnership.v14i2.5169>.
- Innocenti, Perla, Mackenzie Smith, Kevin Ashley, Seamus Ross, Antonella De Robbio, Hans Pfeiffenberger, and John Faundeen. 2011. "Towards a Holistic Approach to Policy Interoperability in Digital Libraries and Digital Repositories." *International Journal of Digital Curation* 6, no. 1 (March 8): 111-124. <https://doi.org/10.2218/ijdc.v6i1.176>.
- Lacey, Sajni. 2019. "Job Precarity, Contract Work, and Self-Care." *Partnership: The Canadian Journal of Library and Information Practice and Research* 14, no. 1. (May 30). <https://doi.org/10.21083/partnership.v14i1.5212>.
- Lambert, Mark. 2019. "Archival Pay and SAA: Some Modest Proposals." *The Southwestern Archivist* 42, no. 2. (May): 5-6. <https://www.southwestarchivists.org/archival-pay-and-saa-some-modest-proposals/>.
- Matienzo, Mark A. 2018. "Evidence of Them: Digitization, Preservation, and Labor." (Presentation, Society of American Archivists, Washington, DC, August 17, 2018). <https://matienzo.org/2018/evidence-of-them-digitization-preservation-and-labor/>.
- Mattson, Rachel. 2016. "Can we Center an Ethic of Care in Audiovisual Archival Practice?." *XFR Collective* (blog), December 2, 2016. <https://xfrcollective.wordpress.com/2016/12/02/can-we-center-an-ethic-of-care-in-audiovisual-archival-practice/>.
- Monahan, Peter. 2019. "Are Temporary Appointments a Threat to Archiving?." *Moving Image and Archive News* (blog), February 8, 2019. <http://www.movingimagearchivenews.org/are-temporary-appointments-a-threat-to-archiving/>.
- Morgan, Paige. 2016. "Not Your DH Teddy-Bear; or, Emotional Labor is Not Going Away." *dh+lib: where humanities and librarianship meet*. (July 29). <https://acrl.ala.org/dh/2016/07/29/not-your-dh-teddy-bear/>.
- Rodriguez, Sandy, Ruth Tillman, Amy Wickner, Stacie Williams, and Emily Drabinski. 2019. "Collective Responsibility: Seeking Equity for Contingent Labor in Libraries, Archives, and Museums." *DLF Working Group on Labor in Digital Libraries, Archives, and Museums*. <https://osf.io/m6gn2/>.
- Weber, Chela Scott, Martha O'Hara Conway, Nicholas Martin, Gioia Stevens, and Brigette Kamsler. 2021. "Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections." Dublin, OH: OCLC Research. <https://doi.org/10.25333/zbh0-a044>.

Facilitating Seamless Access Through Collaborative Workflows, Advocacy, and Communication

Martha Anderson, Max Eckard, Melanie Griffin, Emiko Hastings, Deb Kulczak, Chris Powell, Olga Virakhovskaya, Caitlin Wells¹, and Katrina Windon

Abstract: Intra-institutional collaboration is often a prerequisite to meeting the access needs of users of archives. This paper discusses two different approaches to collaboration at the University of Michigan and the University of Arkansas Libraries, both of which are shaped by organizational structure, staffing, and existing processes and technical choices. Common challenges facing our shared desire to provide a seamless access experience for our users are articulated, among them a tangle of poorly integrated systems, the use of temporary or term-limited staff, and the fragility of collaborative relations when they are largely based on time-bound technology projects or personal relationships rather than organizational structure. Generalizable solutions for approaches to both technology projects and services more generally are suggested for fostering ongoing collaboration within institutions while preserving the separate identities of individual units within them.

Introduction

Intra-institutional collaboration is often a prerequisite to meeting the access needs of users of archives. Yet substantial barriers to collaboration and seamless access — be they resource-based, communication, standards-based, technical, or administrative — abound.

Institutional contexts can often feel unique, and internal divisions may feel distinct, even siloed, to those working within them, but to external stakeholders, they may all be construed as a single entity or confused with each other. Additionally, practitioners may have only a small window into the work of their colleagues, focused at the point that their collaborative work overlaps, and may be unaware of other commitments and projects occurring simultaneously.

This paper discusses two different approaches to such collaboration. One focuses on interactions among administratively separate archival repositories and an operationally separate Library Information Technology division at the University of Michigan, the other on interactions between a single archival repository and other functional units at the University of Arkansas Libraries. In both cases, a variety of platforms and descriptive practices provide information and access to users.

¹ Caitlin Wells left the University of Michigan Library midway through the Lighting the Way Working Meeting, and it is possible the Special Collections Research Center is not completely represented in the University of Michigan case study.

However, behind the scenes, the work of staff is mediated by an even larger ecosystem of systems and technology, and the continual maintenance they require. In discussing each institutional context in detail, we look to articulate common challenges, but more importantly to suggest generalizable solutions for fostering ongoing collaboration within institutions.

Case Study: University of Michigan

Background

The University of Michigan (U-M) case study group includes representatives of three administratively separate archives: [Bentley Historical Library](#), [Special Collections Research Center](#), and the [William L. Clements Library](#), and documents their interactions with the [U-M Library Information Technology \(LIT\)](#) division to share archival collections and their metadata across and through various discovery and delivery platforms.

Special Collections and LIT are both administratively part of the U-M Library; Bentley and Clements are separate libraries on campus. All three archives, however, rely on several services provided by LIT, who manage, design, develop, and maintain the technology environment for the U-M Library. This includes working with libraries, archives, museums, and academic departments across campus, as well as other academic institutions within the state of Michigan.

While all involved aspire to facilitate seamless access through collaborative workflows, advocacy, and communication, the separate institutional approaches to description and digitization, shaped by different collections and institutional histories, have often made collaboration challenging. Disparate levels of technical expertise, staffing, and financial resources, as well as separate administrative structures, contribute to the situation.

Archival Institutions

Bentley promotes the study of the State of Michigan and the University of Michigan. Its holdings fall mainly in the 19th-21st century, with strengths in the state's political and social history, history of the U-M, and architecture. Special Collections' materials are broader in scope, and include radical political and social movements, transportation history, culinary studies, filmmaking, and post-Beat poetry. In addition to archival material, Special Collections also houses the U-M Library's rare books and special collections. Materials collected range from medieval manuscripts to 21st century born-digital material. Clements collects primary source materials related to the Americas, with strengths in 18th and 19th century American history. Materials include rare books, manuscripts, maps, prints and photographs.

These differences in archival collecting scopes, of both time periods and formats, impact choices made in the creation of metadata and the delivery of digitized archival materials, extending to the use of software and workflows, as well as access restrictions to online collections. For example, since Clements' holdings include rare books and relatively small manuscript collections, description is generally quite granular; for Bentley and Special Collections, whose collections can range in size from single folders to hundreds of linear feet or terabytes of data, the level of aggregation for description is more variable.

All three archives are engaged in digitization of their collections, though differences in the nature of collections, formats, and priorities reflect on institutional staffing and workflows. Bentley increasingly digitizes their physical records, including A/V material, via both in-house and vended digitization, and regularly curates born-digital and web archives (archived websites, YouTube channels, archived social media, etc.). Special Collections has also started to collect a wider range of born-digital material in addition to more traditional print and A/V. The unit currently digitizes A/V material through an outside vendor for access purposes and has recently resumed the process of digitizing print material. Clements collects traditional paper formats and digitizes selected collections in-house.

Due to the more contemporary nature of its holdings, a significant number of Bentley collections have access, copyright, and duplication restrictions, according to individual donor gift agreements, official U-M records policies, state and federal laws, and internal policies. In some cases, duplication is prohibited. Access to digital materials can be limited to authenticated members of the U-M campus community or those physically present in the Bentley reading room. Special Collections materials may also have restrictions based on copyright and donor gift agreements. Duplication of material may also be restricted based on staff time and condition of the material. Most Clements collections are in the public domain and do not have access restrictions; some 20th century archival collections may have copyright restrictions, but in general this does not restrict duplication and delivery.

LIT's Central Role in Providing Access

LIT leads or is a key partner in most of the U-M Library's technology initiatives. This includes many of the software applications and technology platforms used by the three archives, both behind the scenes and for online delivery of finding aids and digitized collections to researchers. All three archival units contribute records to the U-M ILS called Library Search, maintained by LIT, which is currently in the process of migrating to Alma.

Access to born-digital and digitized materials and finding aids is provided through DLXS, the software platform LIT developed in the early 2000s. Initially designed and built for access to continuous tone art images and scanned books from U-M Library's collection, it has been expanded and adapted over the years to deliver finding aids, digitized archival materials organized in folders, and some multimedia as well. The Bentley, Special Collections, and Clements finding aids have a shared origin point derived from the Bentley original templates. The common genesis of these EAD profiles makes it possible for the DLXS finding aids component to host three separate EAD collections with shared functionality and architecture and only minimal differences in appearance.

Despite their common origin, a variety of methods are used to produce EAD across the three repositories. Bentley staff have developed a custom EAD exporter for their ArchivesSpace instance to facilitate EAD delivery to LIT. As Special Collections' ArchivesSpace instance has not been customized, minor edits to support desired DLXS interface functionality are done in Oxygen after EAD is exported from ArchivesSpace. Clements uses ArchivesSpace only for accessioning; finding aids are written in Microsoft Word and converted to XML format using Word macros originally developed by the Bentley. The resulting EAD files are edited in XMetaL to clean up any conversion errors. LIT hosts three separate instances of ArchivesSpace to accommodate the three archives' differences in customization and use of the system, particularly when it comes to publishing workflows.

Over time, each of the archival units have requested interface changes for their individual finding aids and digital collections interfaces, and the behavior of various collections has been modified by LIT to search and display unique metadata for the digitized archival materials. Aeon requesting has been separately integrated with DLXS for all three archives, initially with a shared user database hosted by LIT. However, Bentley is now hosting its own server and Special Collections and Clements are moving to separate cloud-hosted servers outside of LIT.

LIT also hosts Deep Blue Documents, U-M's institutional repository, and Dark Blue, dark repository that provides long-term storage for preservation versions of digitized A/V material and medium-term storage for forensic images/file transfers of born-digital archival accessions. Bentley and Special Collections make use of these repositories for born-digital textual materials and web archives as well as A/V material.

Staffing Differences

Description and digitization of archival collections at Bentley occurs within a specialized curation team that takes a format-neutral approach to collections processing, management, digitization, and access. This is a large team of 10-12 members. This approach allows Bentley to operate relatively independently and support a variety of approaches to digitization, both in-house and vended, project-based and on-demand. A critical component of processing and description and sometimes digitization is the use of term-limited Project Archivists that serve term-limited positions of two years. Students also help with a variety of projects that complement all curation processes.

Special Collections has a much smaller staff; there is one Collection Services Librarian who oversees books as well as archives, and one Processing Archivist. As a result, it relies on student workers and other term positions, which means there is a significant amount of staff turnover. In-house digitization capacity is limited and largely focused on filling patron requests.

Clements has four curatorial divisions for Books, Manuscripts, Maps, and Graphics. Both Manuscripts and Graphics Divisions have a curator who accessions and oversees the creation of EAD finding aids and MARC records for their respective archival materials, with the assistance of other library staff, student workers, and volunteers. The Digital Projects Librarian, with one Digitization Technician, is responsible for the digitization of materials from all four divisions.

Current Points of Coordination

On an individual level, staff at the three archival institutions and LIT have multiple points of coordination and collaboration. These connections have been invaluable for the resolution of specific issues in daily operations, including collection development and management, cataloging, finding aids maintenance, and reference and teaching. However, these ties rely on personal connections and institutional memory, which can be easily lost with staff turnover.

The three institutions occasionally coordinate on collection development, either referring potential donors to one of the other repositories or transferring materials when they are a better fit for another institution. Special Collections and Clements have sometimes made joint purchases or traded materials for long-term loan.

All three archival units have collaborated with LIT in various aspects of software assessment and design, primarily around DLXS, but most recently in an investigation of ArcLight and in planning for moving digital collections from DLXS to a new digital repository and access interface. In addition, Bentley, Special Collections, and Clements members have served on a variety of U-M Library committees or served as liaisons on library technology investigation or implementation projects.

Barriers and Challenges

Among the three archival institutions, there is a strong desire to retain local practices that benefit the unique collections and needs of each unit. Despite the common origin of their EAD creation, the institutions have historically been wary of collaboration if it means sacrificing autonomy or giving up locally customized solutions in favor of one-size-fits-all standardization. A solution that works for one unit may not work for another without substantial modification, and the differing policies and priorities of each unit have occasionally made it easier to move ahead separately rather than take the time required to find common ground. With finite time or internal priorities that require action, one institution may feel pressured to “go it alone” in order to get a workable solution completed in a timely manner. In other cases, a shared solution becomes unsustainable when institutions end up going in different directions.

Collaboration is also challenged by the disparate levels of technical expertise, staffing, and financial resources of each unit. Clements and Special Collections have less in-house technical expertise and rely almost exclusively on LIT to provide technical support. Bentley has more technical expertise, but that sometimes results in an expectation that they will figure out the solution and create a process for everyone else to follow. Differences in in-house technical expertise and financial resources can make it hard to align on what kinds of software or support are needed or possible. Meanwhile, LIT understandably does not want to support multiple systems or separate modifications for each partner, which make maintenance and migration to new finding aid and digital collections platforms more difficult. Staff turnover and limited staff capacity also play into this, as it can be especially difficult to find time to collaborate when institutions are already stretched thin and understaffed.

From the LIT perspective, challenges arise in attempting to support the differences in configuration and use of platforms like Aeon and ArchivesSpace when the archival units cannot come to consensus on standard use and display, as is required for the shared Library Search catalog used by all three institutions. Moving off of an aging but essential platform like DLXS is challenging, with so much content and so many customizations for the various digital collections (archival or not) over the years.

Institutional differences are exacerbated by an overall culture of decentralization at the university that in some ways de-incentivizes collaboration. With separate administration and budgets for LIT, Bentley, Special Collections, and Clements, it is sometimes unclear how cost-sharing is supposed to work, or whether the administrative priorities of one unit align with the others. Shared projects are often done on an ad hoc basis, with no official memorandum of understanding between the different parties; this lack of formal agreements endangers grassroot collaborations that remain fragile. The repositories benefit greatly from the free resources and expertise provided by LIT, but do not want to exceed the unspoken boundaries of the partnership by asking for more than is reasonable; that said, learning what is “reasonable” has also taken time.

Impact on Staff and Users

All of this has negative impacts on staff. There are inefficiencies in maintaining parallel systems, and lack of collaboration means that there are few opportunities to learn from or benefit from others' experiences. The use of temporary labor for essential operations contributes to this problem. Having dedicated permanent staff preserves and deepens the institutional knowledge that is so essential for cross-institutional collaboration, while frequent turnover undermines the accumulation of knowledge. In addition, constant re-hiring and re-training occupies a lot of the permanent staff members' time, whose remaining energy is concentrated on meeting the basic primary job responsibilities within their units.

While the use of the same finding aids platform, digital collections, and Aeon request software results in a consistent and familiar user experience for researchers at each of the three institutions, it also causes confusion. The archives share the Library Search catalog, but have three separate finding aids sites and separate digital collections that are not easily cross-searchable. In addition, researchers must register with each library's Aeon system separately and navigate different library hours and reading room policies. People who are doing research on a topic where very similar material is held in multiple archives therefore need to request from and visit multiple archives. This has been confusing for users and often requires in-person remediation.

This confusion is not caused by technology alone. Anecdotally, researchers and donors often confuse the collecting scopes of all three archives and sometimes arrive in person at the wrong one. The current finding aids platform unintentionally strengthens the confusion, because of the similar but siloed user experiences. A future shared finding aids platform, possibly ArcLight, will make it easier for users to find all the collections and search across them, but may not resolve confusion about the separate identities of the institutions.

Future Directions

As outlined above, Bentley, Special Collections, and Clements significantly differ in collections scope and size; staff size, structure, and specialization; and application of technology. However, as we were working on this article, we found important similarities and common challenges that make coordination not only feasible but important. Finding points of coordination while preserving our separate identities could help researchers have a more seamless archival experience at U-M.

Our Lighting the Way experience was a great and rare opportunity to “compare notes,” even though we’ve known each other for years, and even occasionally served on the same committees. Future possibilities include (but are not limited to) establishing regular opportunities for discussions, learning more about each other, and sharing best practices and expertise. These discussions can and should go beyond use of technology, procedures, and workflows to encompass the issues of equitable representation and access, inclusive collecting and description, diversifying our staff, and more. We can also benefit from communicating institutional priorities for each unit, so that others can better understand the context for potential collaborations.

Taking advantage of U-M Library's implementation of Alma in summer 2021 is a logical point of cooperation, as all three institutions contribute to the shared U-M Library ILS, Library Search. It will be important to advocate for the needs of the three archival institutions within the larger library catalog.

U-M Library’s exploration of technology and services to support a new Digital Collections Platform to replace DLXS also represents a needed point of collaboration.

Case Study: University of Arkansas

Background

The University of Arkansas Special Collections was formed in 1967, with a mission to promote research and scholarship of the history, culture, and people of Arkansas and the Ozarks. Since that beginning, it has grown to a staff of 14 and a roster of collections and researchers that stretch far beyond state and regional borders. As our researcher base has expanded and diversified, so too has the need for archival description that can reach them where they are.

Although it serves as the physical and intellectual home of the University’s archival collections, Special Collections is far from alone in working to promote those collections to researchers, and it relies heavily on the expertise of other units within the Libraries. While nearly all Libraries staff share in this work to some degree, we focus here on the three partner units that engage most regularly in work to provide access to archival collections: Content Services (formerly Technical Services), which catalogs rare books and published Arkansiana, ingests MARC records into the Libraries’ ILS and OCLC, and provides metadata for digital projects; Digital Services, which manages all digital exhibits, as well as patron requests for digitization of print materials; and Web Services, which maintains the Libraries’ web presence, and provides support in integrating different access platforms and tools into a unified whole. These partner units have varying staffing levels and have wide-ranging responsibilities beyond their commitments to Special Collections. The Content Services Department has had a dedicated Special Collections MARC cataloging unit since the 1990s. Consisting of one cataloging librarian and one cataloging assistant for much of this time, the unit loosely coordinated its work with the Special Collections Department until 2017, when the two departments began meeting regularly to discuss upcoming projects and set priorities in tandem. At that same time, other Content Services staff working with Special Collections materials—other catalogers and staff from serials and preservation—also joined these meetings. In this way, collaboration was increased, even though only one Content Services staff member catalogs full time for Special Collections.

The collaboration between Special Collections and Content Services on digital projects is similarly longstanding. While the earliest digital projects were solo Special Collections efforts, beginning in 2011 they became joint projects, with archivists (the subject experts) and catalogers both determining the metadata elements to be used and supplying descriptive metadata, while catalogers provided controlled vocabulary and overall quality control. The Libraries’ application profile—the “CONTENTdm Cookbook” (University of Arkansas Libraries, 2021) —was also drafted by Content Services and Special Collections working together. Over the years, the number of Content Services personnel contributing to digital projects has grown from one dedicated cataloger to a group of catalogers and support staff working under the direction of the department head. When the Digital Services Department was created in 2015, we adopted a true team approach to digital projects, with representatives from each department serving on every project team. Additionally, the heads of the three departments or their representatives meet monthly to talk about priorities and project timelines.

In addition to work on digital projects, Digital Services provides crucial research support by imaging and processing Special Collections researcher digitization requests. It also serves other units on campus. Roughly fifty percent of its staff is funded by grants and crowdfunding, so externally funded projects take priority. The unit takes a team-based approach to its work. For a small-scale, one-time request, a team member from Digital Services with the necessary expertise for that project is assigned to the task. For large-scale digital projects, the entire team of about eleven workers is assigned to the project until completion. Digital Services has found this approach very successful in delivering digital requests in a timely manner. In addition, the team-based approach provides workers with greater flexibility and a more diverse work experience.

The Web Services Department, in addition to maintaining various websites, oversees our ILS and customizes interfaces for CONTENTdm, QuickSearch (Summon), Aeon, and ArchivesSpace. Responsibilities are generally distributed between two Web Services personnel, without one person being assigned exclusively to Special Collections projects.

Processes

These different units have, in accordance with their different cultures and missions, evolved somewhat different approaches to planning and overall workflows. Some aspects of Special Collections work, for instance, are heavily driven by an annual planning cycle even as other work of the unit is less predictable and driven by donor and researcher demands. Some units, such as Web Services and Digital Services, tend to be more project management-driven. These different approaches to workflows and scheduling are rife with potential for misunderstandings and frustration when coordinating and scheduling projects among units, which makes frequent communication and shared planning key.

Philosophies

Beyond different approaches to structuring work, these units also have different understandings of the best ways to structure and present information. While most University of Arkansas Libraries faculty have some shared educational background — an ALA-accredited master's is a typical requirement of faculty positions — best practices and standards in their specialized fields may vary widely, and these frameworks in turn shape how they conceive of projects, users, use cases, and description.

Key to the principles of archival description is the idea of aggregate description — that, as DACS puts it, “[d]escription of the aggregate is ... an indispensable component of establishing context and must be provided before proceeding with the description of component parts” (SAA TS-DACS 2021). On the other hand, key to the nature of digital collections is the fact that researchers often arrive at a digital object page without ever viewing the collection landing page or the finding aid for source collections. In many cases, an archival collection that has been fully processed is, when drawn upon for a digital collection, reprocessed in a sense as certain items are given new item-level description. Traditional cataloging generally falls somewhere in the middle of the archival focus on the aggregate and the digital collection focus on the item, with an emphasis on describing single, published items, but not usually at the level of an individual photograph or letter. In the interest of improving access to Special Collections materials, we have adopted a flexible approach to description. For instance, while we normally describe at the single item level for our CONTENTdm collections, a current project aims to digitize whole folders of archival material. We are treating these as aggregates — compound objects with the object level metadata largely taken from the ArchivesSpace finding aid. The page-level metadata reflects only the

information that is unique to an individual item, such as title, extent, and for photographs, subject terms.

Similarly, practitioners may have different metadata standards that feel more natural or appropriate to them, from EAD for archival collections to MARC for catalog records to Dublin Core (DC) for digital collections. Mapping between these standards is always a compromise and necessitates privileging the structure of the target standard. For traditional cataloging, RDA and MARC are the norm, while Special Collections archivists use DACS and EAD for most of their descriptive work. Our digital projects in CONTENTdm combine DC metadata with descriptive principles derived from RDA, DACS, and various DC best practices guides.

Software and Systems

Like many Libraries, the University of Arkansas Libraries has found itself increasingly enmeshed in a variety of technological systems and solutions. These systems now provide the core infrastructure for our archival description and discovery, and indeed increasingly shape the decisions we make about how we describe, promote, and provide access to collections.

Aeon

In August 2020, Special Collections implemented Aeon to manage researcher accounts and collection use. In our implementation, Aeon integrates with the ILS (Sierra), ArchivesSpace, Caiasoft, ILLiad, and CONTENTdm. All Libraries staff can have researcher accounts that allow for requesting collections materials, but currently only full-time Special Collections and Preservation staff have access to the staff client to process those requests. In addition to facilitating collection use at the individual level, Special Collections also uses the Aeon staff client to manage a variety of collaborative workflows, including interlibrary loan scanning requests, offsite storage retrieval requests, patron-driven digitization requests, and large-scale digital projects. Some of these workflows, including ILL and offsite storage requests, rely on APIs to facilitate communication between units and software systems. Others, most notably digitization workflows, require using email templates built into Aeon to communicate with partner units.

ArchivesSpace

Special Collections uses ArchivesSpace as its archival content management system. Accession records, resource records (finding aids), location information, and donor information are all stored within the staff client, and finding aids are displayed to the public through the public user interface. While data from ArchivesSpace feeds into Aeon (through the Aeon-ArchivesSpace client add-on); our off-site storage inventory management system, Caiasoft (through a workflow involving SQL queries and spreadsheet upload); Sierra (through MARC export); and, in some cases, CONTENTdm (when finding aid data is reused for digital object metadata), the information interchange is typically one-way and mediated by Special Collections, as the only personnel outside Special Collections to have ArchivesSpace client user accounts are Web Services staff. A pilot project is underway to add digital objects to resource records that reference digital objects in CONTENTdm, and plans are in place to implement the ArchivesSpace/Alma integration plugin with the Libraries complete their migration from Sierra to Alma in 2022.

CaiaSoft

CaiaSoft is the storage management system used for materials at the Libraries' off-site storage facility, LINX. It manages location information and circulation but is not the system of record for any item metadata. Archival metadata is derived from ArchivesSpace, and metadata for cataloged works is imported from Sierra. While CaiaSoft is invisible to our end users, it provides tracking and support for the regular transfer of materials between facilities. All Special Collections staff have CaiaSoft accounts, although for many functions, staff commonly interact with it through Aeon workflows rather than directly.

CONTENTdm

CONTENTdm, the Libraries' digital collections display platform, provides access to digitized rare materials. Some of the services provided within the digital collections include access to digital files, robust metadata, full-text searching, downloading, and printing capabilities.

Selected digital collections in CONTENTdm have an Aeon plugin enabled that facilitates direct requesting of high-resolution digital copies without requiring users to switch systems. Staff time is still required, however, to process the request, download a high-resolution scan, and deliver it to the researcher.

Sierra/Summon

The University Libraries implemented its Innovative Interfaces system – currently Sierra – in 1993. The catalog holds records for most of the books, serials, media, and manuscript collections housed in Special Collections (the latter with a link to the online finding aid). In 2016, the Libraries added a Summon discovery layer to the catalog, which can additionally integrate results from our CONTENTdm digital collections. In the current environment, it is not possible to pull in metadata directly from ArchivesSpace.

Systems Access

While the personnel in Special Collections, Digital Services, Content Services, and Web Services all have their various “home” systems, to collaborate effectively they must be familiar with the other systems and standards in play. Most if not all of our Libraries' systems follow the information security principle that access should be limited to those who need it, and that permissions should, when possible, be specific to use cases. This limits risk of confidential information being shared inappropriately, or of records being edited or deleted inadvertently, but it can also create information asymmetries between personnel and units and support perceptions of gatekeeping. As important as the communication between our systems is the communication between our units' personnel to ensure shared understanding of policies and platforms.

How It All Fits Together

Coordinating between all the people, processes, standards, styles, systems, and schedules is not seamless! Yet the more those seams show to the user, the more challenging the information discovery process is likely to be. Currently, the University of Arkansas Libraries are implementing a number of approaches to facilitate our interdepartmental collaborations.

API-driven coordination

APIs facilitate a growing number of collaborative workflows between Special Collections and other units. Due to staffing constraints, the University of Arkansas only uses existing API integrations that either have been developed by the software creator or are openly available. A sampling of API-driven project management illustrates the scope of collaborative workflows possible:

- The ILLiad API allows Interlibrary Loan to send scanning requests to Aeon. Special Collections staff then process, fill, and send these requests to ILLiad via the Aeon staff client;
- The ArchivesSpace API enables an Aeon staff client add-on, which allows staff to search the ArchivesSpace staff interface for container location inside of the Aeon staff client;
- The CONTENTdm and Aeon APIs allow patrons to request access to either the physical object or staff-mediated high resolution scans directly from the Digital Collections discovery system;
- The Caiasoft and Aeon APIs allow Special Collections staff to send circulation requests to Caiasoft for retrieval from offsite storage and for offsite storage staff to fulfill the request in Caiasoft.

The benefits of system-mediated collaborative workflows are numerous, including allowing staff to work in their “home” system. In addition, researchers can interact with collection materials at the point of discovery rather than having to navigate to a different interface. Despite the numerous benefits, there are also significant barriers to an API-driven collaborative approach. In the case of the University of Arkansas, we are limited by our reliance on existing scripts and limited capacity to modify those scripts. These limitations would likely be experienced by other similarly staffed and resourced institutions. In addition, even robust technological integrations are not a substitute for staff communication and the shared understanding of workflows that comes from working together.

Automated communication to enhance workflows

When APIs do not exist or do not fit our use case, we often rely on system-generated communication to enhance or streamline workflows. Digital Services and Special Collections, for example, created a systematic workflow for one-time digitization requests to expedite delivery and avoid creating multiple digital surrogates of the same item for different researchers. The workflow culminates in a published digital collection that allows researchers to access digital files from previous one-time requests when right restrictions allow. The workflow requires multiple points of coordination: transaction initiation, process and delivery, and display and preservation. Special Collections staff use Aeon to send a template-based email to Digital Services staff alerting them of new one-time scanning requests. This email includes the item’s existing metadata, generated from the finding aid, and Digital Services staff in turn use a simple crosswalk to map those metadata fields to Dublin Core in the CONTENTdm record. The Aeon email template also prompts Special Collections staff to note if the digital files can be made publicly available or not. During the process and delivery stage, a Digital Services worker is assigned to the selected transaction, and once the work is completed, Special Collections receives an email prompting them to review the digital work, approve the work or request revisions, and retrieve the physical materials from Digital Services. The files from both publicly published and unpublished items are then added to the archival information package of that specific year, including the preservation and

access files for those materials, the related metadata created by Content Services catalogers, and preservation notes if needed.

Sierra, the University of Arkansas' current ILS, does not support a fully automated ingest of collection-level MARC records generated from Arkansas. Instead, Special Collections staff export a MARCXML record from ArchivesSpace, add an 856 field with the finding aid URL to the XML file, and upload the file into a Box folder shared between Special Collections and Content Services staff. Content Services uses Oxygen along with an XSLT stylesheet (Buza 2015) to merge the individual XML files into one. After that, staff turn to MarcEdit for converting the file to MARC21 and performing batch editing. From there, the records can be loaded into our Sierra system and OCLC.

Future Directions

For all of our considerable efforts at improving our workflows and platforms, real user experience issues remain. Depending on the point of entry, users may encounter differing levels of description and availability of access to digitized content. They may need to employ new search strategies and learn new field names as they follow links across systems. There is currently no true single search across all our platforms – not even Google, even though many of our users may expect it to be. Our systems often engage in one-way communication that doesn't incorporate feedback into the source system. Better integrations that rely on available APIs might help us bridge some of these gaps so that metadata is more bidirectional and requires less staff mediation.

We also hope to incorporate more mechanisms for user feedback. Our Libraries are increasingly concerned with assessment and user experience, but no usability testing has occurred at the local level for most if not all of our platforms for archival discovery. Particularly as we begin to think more intentionally about the accessibility and impact of our archival description, we hope to pursue additional avenues for user feedback, drawing on examples like the “suggest a correction” and “ask a question” dialogue boxes in Princeton University's finding aid display system.

It's not just researchers who need more space in our systems and workflows – it's also our colleagues in other Libraries or University units who may have interest, expertise, and valuable new perspectives, even if their job duties don't align with our existing staffing models for archival discovery work. Several recent digital projects have successfully integrated subject librarians as subject selectors (curators) of digital collections. There have been several project CERES awards for which the agriculture subject selector collaborated as a principal investigator and subject matter expert (USAIN, undated). Past collaborations include the *Colonial Arkansas Post Ancestry* and the *Ozark Folksong* digital collections, for which a French language professor and the performing arts librarian contributed as subject selectors.

The Shared Mission

For all the shared systems, shared workflows, and shared frustrations, what really binds together our units in this work is a shared mission – a genuine desire to increase access to archival collections and to serve our students, faculty, and community. This mission is shared between our units at the University of Arkansas Libraries. It is also shared across archival institutions, as documented in writings on collaborations between units at other libraries forging and refining cross-departmental partnerships in archival cataloging (Sweetser and Orchard 2019; Turner and Schuster 2019) and digital projects (Gueguen and Hanlon 2009; Hunter, Legg, and Oehlerts 2010; Perrin and Weaver 2020), and as

experienced in our discussions with colleagues at the University of Michigan as we participated in the Lighting the Way Working Group meetings.

Reflecting on our experiences with cross-division collaboration, particularly systems-driven collaboration, suggests directions for future work in our own organization and the profession more generally. Recognizing and understanding different professional practices across units is key, as is thinking about ways that those differences might work in tandem. Recognizing differences does not, however, imply that the end goal is eradicating those differences. Instead, it asks all project participants to be aware of and respectful of differences, which might range from budget priorities to staffing to descriptive standards.

Such recognition should also drive future use of systems. In our case, an ILS migration from Sierra to Alma/Primo and the implementation of a hosted digital preservation platform in Special Collections are key examples. The ILS migration provides opportunities for re-thinking how all of our systems work together, both at the level of technology and at the human levels. In the case of digital preservation, acknowledging different needs has led to adopting different solutions in the Libraries, one to meet the needs of Digital Services and one to meet the needs of Special Collections. An important component to this work has been remembering that, while our technological needs are different, our core goals — access and use — are shared.

Conclusion

Despite the differences between and within our institutional contexts, several common themes are apparent. A series of common challenges faces our shared desire to provide a seamless access experience for our users through collaboration. We are faced with a tangle of poorly integrated systems, some of which are simultaneously brittle and also central to our work. The use of temporary or term-limited staff means that we are often in survival mode, trying to “make do” rather than “make better.” It also means that collaborative relationships can be fragile when they are largely based on personal relationships rather than organizational structure. And in some cases, eliding the “seams” or differences between systems and institutions may result in confusion for our users and improper resource allocation.

Technology projects — whether adapting existing systems to local use cases through customizations and plugins or migrating to new ones — can offer a fruitful opportunity for collaboration. However, the collaborative relationships generated by these projects are often time-bound and may disappear once the project is complete. Any given technology, while of central importance to the work of archivists to provide access to archives, is also inherently more ephemeral than the content and description held within archives or the people that donate to archives, work at them, or use them; without attention to those relationships, technology projects may only exacerbate existing divisions or create new ones. Stronger ongoing collaborative relationships can potentially be fostered by formal service agreements, particularly with technology-focused units.

Coordination on services ranging from instruction to collection development provides another way of fostering programmatic collaboration based on ongoing operations rather than one-off projects. Team-based approaches which do not seek to erase differences but rather build connections across and between areas of expertise show a great deal of promise, and generating a productive collaboration in one area can often lead to new collaborations in others. Sustained, meaningful alignment on mission,

goals, and policy are crucial to fostering collaborative relationships, whether between or within institutions.

Above all, throughout this process we have realized that we share a number of common goals, primary among them the desire to improve our users' experience and their access to collections by closing the gaps between and among our systems, processes, and colleagues.

References

- Buza, Adrian. 2015. "Merge Multiple XML Files into One XML File," Oxygen XML Forum post, March 26, 2015. <https://www.oxygenxml.com/forum/topic11375.html>.
- Gueguen, Gretchen, and Ann M. Hanlon. 2009 "A Collaborative Workflow for the Digitization of Unique Materials." *Journal of Academic Librarianship* 35, no. 5 (September): 468–74. <https://doi.org/10.1016/j.acalib.2009.06.001>.
- Hunter, Nancy Chaffin, Kathleen Legg, and Beth Oehlerts. 2010. "Two Librarians, an Archivist, and 13,000 Images: Collaborating to Build a Digital Collection." *Library Quarterly* 80, no. 1 (January): 81–103. <https://doi.org/10.1086/648464>.
- Perrin, Joy M., and Robert G. Weaver. 202. "Context Is Key: Library and Archive Collaboration for Digital Projects." *Collaborative Librarianship* 12, no. 1 (January): 68–79. <https://digitalcommons.du.edu/collaborativelibrarianship/vol12/iss1/7>.
- Society of American Archivists' Technical Subcommittee on Describing Archives: A Content Standard (SAA TS-DACS). 2021. "Principles of Archival Description." *Describing Archives: A Content Standard*. Version 2021.0.0.1. https://saa-ts-dacs.github.io/dacs/04_statement_of_principles.html#describing-aggregations-of-records.
- Sweetser, Michelle, and Alexandra A. A. Orchard. 2019. "Are We Coming Together? The Archival Descriptive Landscape and the Roles of Archivist and Cataloger." *American Archivist* 82, no. 2 (Fall/Winter): 331–80. <https://doi.org/10.17723/aarc-82-02-18>.
- Turner, Rachel Berman, and David Schuster. 2019. "Carts in the Hallway: Cataloging and Special Collections—A Partnership for Success." *Technical Services Quarterly* 36, no. 2 (April): 142–54. <https://doi.org/10.1080/07317131.2019.1584982>.
- University of Arkansas Libraries. 2021. "Staff Guide: CONTENTdm Cookbook." <https://uark.libguides.com/contentdmcookbook>.
- USAIN. "Project CERES Goal Is Preservation." Website of USAIN. Accessed June 29, 2021. <https://usain.org/Project-CERES>.

Section 2: Assessing and Applying Standards and Best Practices

Lost Without Context: Representing Relationships between Archival Materials in the Digital Environment

Jodi Allison-Bunnell, Maureen Cresci Callahan, Gretchen Gueguen, John Kunze, Krystyna K. Matusiak, and Gregory Wiedeman

Abstract: The problem of representing context for archival materials in digital asset management systems (DAMS) has been noted - and lamented - for as long as digital representations of archives have been online. This white paper discusses the nature of this challenge, explores why it remains so thorny, and provides examples of where archival access systems have been successful in representing context. With hopes of moving the conversation forward, we provide a set of principles for representing archives in context that can be implemented regardless of the particular systems employed. These principles are based on archival standards and software best practices, and can be summarized as six ideas:

1. Create space for deep conversations with all stakeholders and so that everyone understands foundational requirements.
2. Value archival context and design systems so that contextual relationships between records are explicit and clear.
3. Leverage the power (and cost savings) of aggregate digitization and description when appropriate.
4. Be consistent about modelling relationships between an analog object (if relevant), a digital object, and the description of the archival record.
5. Use persistent identifiers.
6. Lean on widely-used standards, systems, and solutions.

Finally, we call on standards-making bodies to introduce a more robust data model for archival representation that includes both the description of archival contents and contexts.

Introduction

In archives, everything comes from somewhere. A postcard in a collection could be part of a body of correspondence, could have been found pasted into a scrapbook, or could have come from a creator's subject file about a particular place. The archival object (the postcard) is described in a finding aid that may be surrounded by widely varying other materials — and so may be titled very differently depending on circumstances. Whether the file of materials is called “Letters from family 1942-1960”, “College

scrapbook 1950”, or “Subject file - Yellowstone”, the postcard is part of that file, which represents archival context that is critical to understanding the object.

This task — representing the experience of understanding materials in the context of how they were produced and used — has been a central challenge for the representation of archival materials in digital asset management systems (DAMS) since the first days of web display.

When presenting digitized materials from archival collections, archives face a variety of risks: lost archival context, unstable digital object links, and degraded user experience. Context is lost when archival objects are imported into DAMS with data object models that do not account for context, and cannot act upon it or reflect it back to users. Physical access is replaced by web links (URLs) that are often unstable. Despite persistent identifiers (PIDs) being a best practice for over twenty years, stable links are still missing from important DAMS.

That user experience is further degraded by missing or duplicative metadata. Archival description relies on aggregate metadata that has been applied to containers and is part of the archival context. Item-level description is often cost-prohibitive, but it can also remove needed context. Without such metadata, it often happens that downstream systems (notably aggregators) end up with hundreds of distinct items that suffer from both identical metadata and missing archival context. Digitized archival materials is presented without context and with a loss of important historical evidence.

Here, we will outline the problem and provide recommendations for DAMS creators and implementers so that all forms of evidence — content, context, and other administrative interventions — can be maintained and understood. The ultimate goal is to improve user experience in understanding the context of archival documents and to support meaningful archival research.

The Problem at Hand

Digitization has offered opportunities for expanding discovery and delivery of archival collections. This has been an enormously important development for democratizing access to archival records. Researchers no longer have to travel long distances and schedule visits during working hours to have access to evidence of the past. At the same time, as archives have used systems designed for other domains that did not consider archival theory and practice, archivists and researchers have faced new challenges for archival representation and end-user understanding of digitized documents.

In the online environment, users often lack the contextual information and interpretative framework that are critical to understanding archival documents and to sense-making in archival research. Archival documents are unique information resources that gain meaning when presented with the associated provenance and background information and in the context of other documents in the collection. Understanding archival content is a complex interpretive and associative process that “requires the performative creation of meaning in relation to material records” (Duff, Monks-Leeson, and Galey 2012, 70). However, it remains an open question how original order can be applied in the digital library environment where there can be more than one way to access archival records (Trace 2020, 341-342; Zhang 2012, 167).

Context is a unifying principle of archival representation (Nesmith 2005, 259-261; Yakel 2003, 22). The contexts of recordkeeping — how historical materials were created and how they relate to one another — is important evidence that is often used by researchers to understand the historical process by which materials were created, used, exchanged, and modified over time. Maintaining the integrity of historical evidence is a core value of archivists' work. However, in the fluid and malleable digital library environment, objects are often separated from original collections and devoid of meaning that is conveyed in the multi-level hierarchical structure of archival description. A recent study of forty-two digitized archival collections found that metadata records fail to indicate the original context of digital surrogates (Force and Smith 2021, 102-104). The authors state, “contextual information about these digital surrogates, such as their provenance, is mostly absent, thereby potentially obscuring their true evidentiary value” (Force and Smith 2021, 104).

The lack of attention to contextualization, representation, and use of digital archives was noted almost two decades ago, when Margaret Hedstrom wrote her paper on “interfaces with the past” (Hedstrom 2002, 23). The “interface” is a metaphor for archivists interacting with users, but in the digital library world, most archivists have little or no control over the interface. The problem is even more urgent with the trend towards more minimal archival description and calls for large-scale digitization of archival collections (Greene and Meissner 2005, 236-249; Miller 2013, 527-533).

Many DAMS employed to display archival materials are based on a data model and metadata schemas traditionally used in library cataloging practices, a framework that does not incorporate network structures currently used to represent archival context. This model is implemented in the older generation of DAMS, such as CONTENTdm as well as newer open source systems like Omeka or Hyrax. The bibliographic data model that assumes that materials can exist and be interpreted as sole items is dominant. The bibliographic model also requires more granular item-level description. Cal Lee demonstrates, calling on archival theory, that making meaningful use and sense of digital objects requires multi-faceted contextual information (Lee 2011, 106) to be meaningfully understood.

More recent systems have been developed with archivists as substantial stakeholders, including the ArchivesSpace public interface and ArCLight. However, those systems remain institution-specific (ArchivesSpace, for instance, was never designed for cross-institutional multi-tenancy) or are not available to institutions seeking to provide combined access to library and archival materials together. Since libraries, archives, and museums utilize vastly different data structures and descriptive controls, it is challenging to use systems from each others' disciplines without fundamental modifications. So much time and energy can be spent trying to define and describe the issue and to capture all possible permutations of it, that often little is left over for creating solutions.

The challenges of archival representation in digital libraries are compounded in the distributed large-scale systems, such as the Digital Public Library of America (DPLA) or Europeana, which harvest metadata from individual libraries or regional aggregators. Metadata records shared with aggregators often do not include links to original collections or finding aids, so results are returned with little relevant contextual information that would help users understand the documents and see the relationship to other records (DPLA Archival Description Working Group 2016, 19-27). Users must navigate two- or three-step pathways to locate digital objects and metadata records at originating institutions and sometimes get lost in the multi-layered structures (Matusiak 2017, 165-167).

All of these challenges are additionally compounded by lack of resources. Time, money, and staffing are all relatively scarce in archives, and little can be spared to build specialized systems for digital discovery, let alone convince our colleagues to fundamentally rethink their own so that we can participate. Additionally, there is no one technological solution that can be adopted by all. There are many ways that context can be conveyed with content. When our few resources are spent designing and implementing differing systems, we may not end up with complementary approaches.

Two examples illustrate both the technical and intellectual problems presented in the current environment.

Example 1: An item contains adequate contextual information in its original system, but be stripped of said context when reused in other systems without adequate infrastructure.

An item described as "Certificate from the French Gallery" from the Philadelphia Museum of Art illustrates this nicely. The Museum's website displays this individual item with reference and links to the collection it comes from, the John G. Johnson Archives.

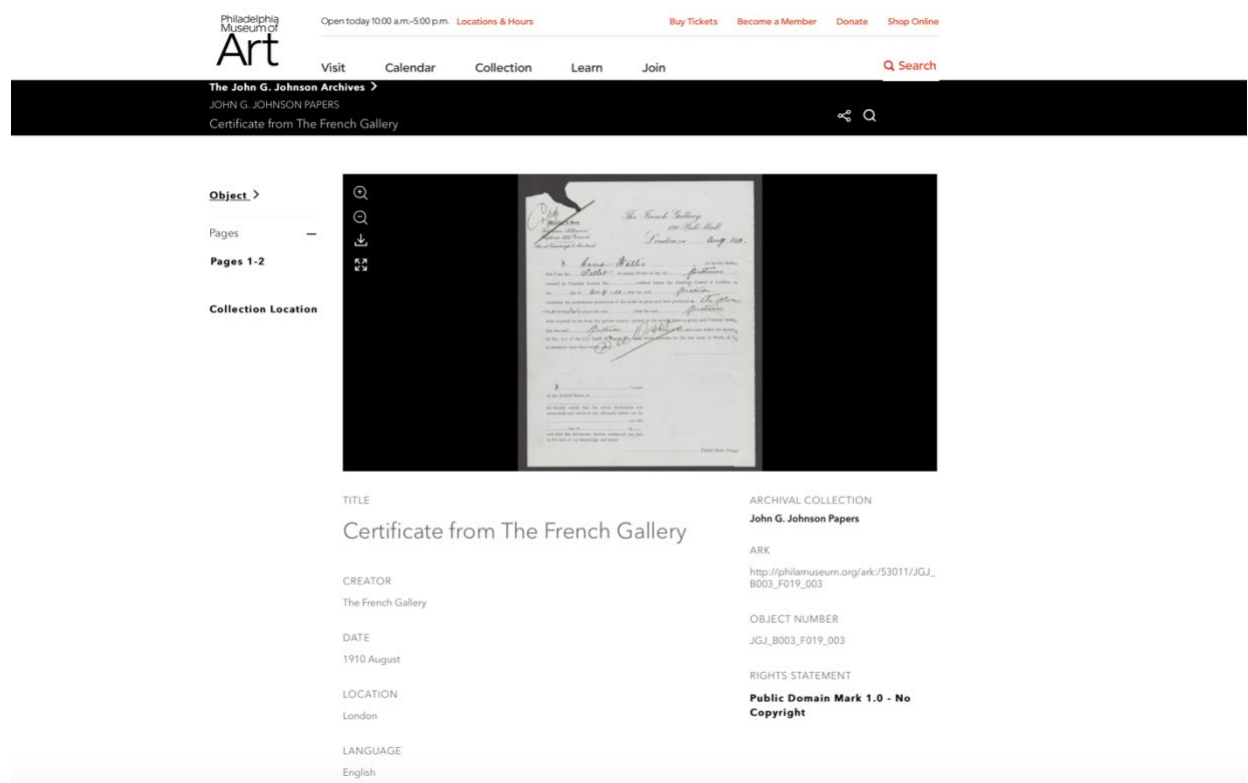


Figure 1: "Certificate from the French Gallery" in original DAMS
(https://archives.philamuseum.org/jgj/JGJ_B003_F019_003, accessed 2021 August 30)

Within that collection, this item is identifiable as part of Johnson's correspondence related to his acquisition of artworks that make up part of the museum's collection. Johnson is not well known,

however, the notes included in the full collection description (not pictured) give adequate context to understand his role in the museum's history.

The same object appears in the Digital Public Library of America, again named "Certificate from the French Gallery." The visible metadata on the item page mentions the collection and the Philadelphia Museum of Art. However, the link to the finding aid and ARK identifier are both part of the full DPLA record that can only be viewed through the DPLA API – not the web-page view that most users will interact with.

All items > Certificate from The French Gallery
Please read [DPLA's Statement on Potentially Harmful Content](#).

Certificate from The French Gallery



View Full Item

↗



Cite this item

Add to a new list

Created Date	1910 August
Description	Certificate from The French Gallery.
Creator	The French Gallery

Partner	PA Digital
Contributing Institution	Philadelphia Museum of Art
Collection	John G. Johnson Papers
Location	London (England)
Type	text
Format	Certificates
Language	English
URL	https://archives.philamuseum.org/jgj/JGJ_B003_F019_003 ↗

Figure 2: "Certificate from the French Gallery" in DPLA
 (<https://dp.la/item/3d29761d3cad51c5838230acfeba360>, accessed 2021 August 30)

The other metadata is exactly the same as it is on the original page, but without the context it is less useful. The collection name and a link to the finding aid are both part of the actual DPLA record, which can be viewed through the DPLA API, but both fields are suppressed from the visible web-page view that most users will interact with. Moreover, the ARK persistent identifier has been suppressed in display (even though it is present in the harvested metadata), while the less stable URL has been carried forward.

DPLA's partnership with Wikimedia Commons, through which DPLA shares its metadata with Wikimedia, shows how the problem quickly compounds. In the new instance of the French Gallery certificate, the collection name and description have been completely stripped from the item leaving it with no contextual metadata save a link (but not the persistent link) to the item in its original context buried near the bottom in a list of links.

File:Certificate from The French Gallery - DPLA - 3d29761d3cad51c5838230acfeba360.jpg

From Wikimedia Commons, the free media repository

Download all sizes
Use this file on the web
Use this file on a wiki
Email a link to this file
Information about reusing

Size of this preview: 485 × 599 pixels. Other resolutions: 194 × 240 pixels | 389 × 480 pixels | 486 × 600 pixels | 800 × 988 pixels.
Original file (800 × 988 pixels, file size: 162 KB, MIME type: image/jpeg)

Open in Media Viewer

File information Structured data

Captions Edit

English Add a one-line explanation of what this file represents

Summary [edit]

Certificate from The French Gallery (info)

Creator: The French Gallery
Title: Certificate from The French Gallery
Description: Certificate from The French Gallery.
Date: 1910 August
Collection: Philadelphia Museum of Art

Source/Photographer

This file was contributed to Wikimedia Commons by Philadelphia Museum of Art as part of a cooperation project. The donation was facilitated by the Digital Public Library of America, via its partner PA Digital.

- Philadelphia Museum of Art identifier: pdigPMA-JGJ_B003_F019_003
- Source record: https://archives.philamuseum.org/jgJ_B003_F019_003
- DPLA identifier: 3d29761d3cad51c5838230acfeba360

Permission (Reusing this file)

This media file is in the public domain in the United States. This applies to U.S. works where the copyright has expired, often because its first publication occurred prior to January 1, 1926, and if not then due to lack of notice or renewal. See this page for further explanation.

This image might not be in the public domain outside of the United States; this especially applies in the countries and areas that do not apply the rule of the shorter term for US works, such as Canada, Mainland China (not Hong Kong or Macao), Germany, Mexico, and Switzerland. The creator and year of publication are essential information and must be provided. See Wikipedia:Public domain and Wikipedia:Copyrights for more details.

w:Wikipedia:Public domain

File history

Click on a date/time to view the file as it appeared at that time.

	Date/Time	Thumbnail	Dimensions	User	Comment
current	22:10, 14 Apr 2021		800 × 988 (162 KB)	DPLA bot (talk contribs)	Uploading DPLA ID 3d29761d3cad51c5838230acfeba360

You cannot overwrite this file.

File usage on Commons

The following page uses this file:

- File:Certificate from The French Gallery - DPLA - 3d29761d3cad51c5838230acfeba360.jpg

Hidden categories: Media contributed by the Digital Public Library of America | Media contributed by PA Digital | Media contributed by Philadelphia Museum of Art | PD US | Artworks without Wikidata item | Files with no machine-readable author

Figure 3: “Certificate from the French Gallery” in Wikimedia Commons
(https://commons.wikimedia.org/wiki/File:Certificate_from_The_French_Gallery_-_DPLA_-_3d29761d3cad51c5838230acfeba360.jpg, accessed 2021 August 30)

This example illustrates how vital contextual information can be lost over time by sharing works between systems that are not designed to capture such information. Although the original source has adequate context, once the metadata is adapted to new systems it is lost.

Example 2: An item in a DAMS exists within an archival collection, but the item's metadata does not acknowledge this.

The following example is simplified from real materials in real archival collections. Three objects exist in a digital library repository. Each has been labeled, by its creators, “communist propaganda.” Looking at each object, the researcher sees the exact same materials — newspaper clippings from 1949 from the *Daily Worker* (the newspaper of the Communist Party of the United States) about the Smith Act trial of eleven Communist Party leaders. But these are indeed from three distinct archival collections — one of these is from the records of Judge Harold Medina, who adjudicated the trials, one is from the records of the American Civil Liberties Union, and one is from the records of the Communist Party, USA. Three copies of the exact same file — three digital objects — three distinct archival contexts.

However, in typical DAMS, these distinct contexts are de-emphasized. The user will have access to information about the clipping itself, but much less information about who was the collector, where it exists within a collection, and how it came to be there. After records are exported from archival systems, there are usually no clues to be gained from the interrelation of records, the intuited recordkeeping practices of creators, or the extra-textual evidence about chain of custody, appraisal, or other intervention that is available in a finding aid.

Context within a single collection is important, too, for historical understanding. Was this file from Judge Medina's own subject files, thereby betraying a prejudice he may have had about the nature of the Party's communications? Or was it from his file of public response mail — a letter-writer may have sent this to him as an example of “Communist Propaganda.” In the Communist Party's records, was this meant as a working file, a place where members could go to get examples of communications (an affirmation that propaganda is desirable and should be produced to help change public opinion)? Or was it stored in its files about the trial, a tongue-in-cheek way to think about its response? In the ACLU's records, are these from the files of a case that the ACLU was invited to write an amicus brief about? Or was it from a central library subject file, so that the organization could reference information about changing political landscapes? In any of these scenarios, contexts of creatorship and the institutional use of records matter very greatly for any researcher who wants to understand how individuals and organizations encountered these ideas and participated in conflicts.

In most DAMS, the information found in item-level records must stand for itself. While not all users or objects require the missing context, by leaving this information behind, DAMS are only fulfilling a portion of the potential users and use cases.

Principles for Access with Context

Principles of archival representation in the context of digital archives that should be considered by decision makers as they choose and develop systems for meaningful representation of archival materials are proposed below. The aim of these principles is to preserve contexts across records, the relationships to records creators, and the events that affect how records are understood over time.

Create space for deep conversations with all stakeholders so that everyone understands foundational requirements.

Sometimes the most important conversations are the ones that seem most obvious. It may not be common for practitioners who come to a digital archives project from the library world, software development, archival administration and museology to have the same understanding of the nature of what is being represented and what a system needs to be able to do. But the development of shared understanding is critical, and one in which respect for both shared and distinct expertise is essential to success, since an element of the persistence of this problem is valuing one area of expertise over another. This is an opportunity for participants to come to consensus — and compromise — about questions of representing provenance, administrative interventions, materiality, and content. These discussions should be concrete and explicit. It may also be helpful to use rapid prototyping methods to compare everyone's expectations of a final outcome to what is eventually — after much time and labor has been spent — produced (Ellis and Callahan 2012).

Value archival context and design systems with context in mind.

In order to preserve context, it's first important to value and understand that contexts across and between groups of records provide essential historical evidence to researchers, and should be preserved whenever possible.

Describing Archives: A Content Standard (DACS) provides a body of rules and principles for creating archival descriptions agnostic of output format. DACS provides useful guidance regarding the essential nature of context:

Within systems that communicate archival description to users, it is often the case that descriptive elements may be shared, inherited, or otherwise linked across and between entities. Traditionally, inheritance has been implicitly presented as hierarchy within the idiom of the print finding aid where frontmatter (collection-level descriptive notes, creator elements, conditions governing access and use, repository information, etc.) applies to archival descriptions on subsequent pages. However, in modern networked archival information systems (relational databases, linked data systems, etc.) linkages, relationships, and inheritances can be non-hierarchical. This makes it particularly important for outputs from these systems to clearly explain relationships so that a user understands which records, agents, or activities an archival description governs. (SAA TS-DACS 2021)

As stated by DACS, representing context is just as crucial for online access to archives as explaining content, however, it may need to be addressed differently than in traditional formats. Attention to maintaining context must be a part of any system design.

One already-existing system designed to maintain context¹ is the EAD-encoded archival finding aid. Archivists providing descriptions in a finding aid benefit from the meanings interpreted from networks,

¹ It is important to note that the EAD-encoded finding aid, displayed as flat HTML, does not entirely successfully nor explicitly explain the relationships between records, creators, and activities. Relationships between aggregations of records are represented by their context within hierarchical XML structures. Ideally, an archivist would explain more explicitly the nature of a single letter to a set of correspondence, for example. The relationships between the letter, the group, and the collection would be explicitly encoded. This, in turn, would

which allows for labor saving. It is not necessary, for instance, when describing many instances of a corporation's board minutes to explain the nature of the company, the composition of the board, etc. Instead, because this has been described elsewhere in the network of archival records the researcher can apply this context to the item and the archivist can simply provide any contextual information that might apply to a particular record (date, extent, etc.). In this way, archival descriptions — as a whole network of related records — become more than the sum of their parts.

Make contextual relationships between records explicit and clear.

DACS compels archivists to not only describe records as information objects, but also to describe the relationships among records, agents, and activities essential to understanding archives (Technical Subcommittee on Describing Archives 2021). In the EAD example given above, context and relationships are conveyed within the network structure of the finding aid. In other systems contextual relationships may be made explicit through the use of additional contextual metadata elements such as collection names and descriptions, or the use of persistent identifiers between systems.

Leverage the power (and cost savings) of aggregate digitization and existing aggregate description when appropriate.

In a system that does not support a networked structure, as described above, the archivist or digital librarian may feel compelled instead to add more information to the item level record so that information found elsewhere in the finding aid may be brought to each individual object. Unfortunately, the process of item-level metadata creation in archives has been widely shown, across cost analyses of digitization projects and programs, to be the greatest cost of the creation of digital archives, so prohibitively expensive as to make digitization at the scale of our users' expectations for online access impossible (DeRidder, Presnell, and Walker 2012, 155-158; Force and Smith 2021, 102-104).

Instead of digitizing and describing a single item within an archives, a digitization project might instead look to see how this item existed in the context of the finding aid and mimic that organization and description. The finding aid provides networked nodes of information objects that existed within the context of one another when they were used in the course of daily life. The description and arrangement archivist, when processing the collection, usually kept these single items together and described them as an entire file.²

In modern DAMS, it is possible to display multiple objects as part of a single description, maintaining in addition the context of the file group. Doing so maintains the arrangement work already done by the archivist who had processed the collection and can minimize additional metadata work. It also creates a level of description appropriate to understanding the material without unnecessary additional information. All digitization will eventually run into limits in funding, labor, and/or time before they reach limits in collections material. By adding description to individual items, archivists are spending

make it possible for an information system to help the user understand the relationships between nodes in a network.

² This is in accordance with DACS principles 9 and 10. "Statement of Principles," Describing Archives: A Content Standard (Version 2021.0.0.2). https://saa-ts-dacs.github.io/dacs/04_statement_of_principles.html#9-each-collection-within-a-repository-must-have-an-archival-description

more time describing materials and serving fewer users than they otherwise could; it is a more ethically sound choice to leverage aggregate description.

Be consistent about modelling relationships between an analog object (if relevant), a digital object, and the description of the archival record.

Thinking of digital objects as aggregates with a single description consistently across systems has the potential to simplify how we think about digital representations.

Within archival management systems and standards, we are encouraged to think of digital objects as analogous to what is described. If an item is listed in the contents of the folder in the finding aid, then we expect to see that item in the digitized folder. Thinking of materials in this way then encourages us to digitize according to the level of description (a box, a folder, etc.) rather than selecting individual items to describe and highlight. Applying this intellectual control consistently across systems removes confusions when linking between them, for example, links in a finding aid to the representations of digital objects in a DAMS.

In ArchivesSpace, Archivists' Toolkit, and other archival management systems, the digital object record can be thought of as an analogue to the container record. Instead of pointing us to where we can find the physical manifestation of what is being described (in a box, on a shelf, in a building), we provide access to the digital manifestation in a networked location. These links can both provide flexibility in systems design, and the potential for describing meaningful aggregations where pointing to each individual item is impractical.

Use persistent identifiers.

One method for retaining context in the representation and reuse of digital objects is through the use of persistent links — in essence, a URL that is maintained by its owner so that it does not break. Even for non-public-facing objects, stable identifiers are necessary for reducing ambiguity and making sure that relationships can be understood and made actionable by machines. Both the object and the description of the object need stable identifiers (DACS 2.1). Despite persistent identifiers (PIDs) being an accepted Internet best practice for over twenty years, stable URL links are still missing from important DAMS (CONTENTdm, PastPerfect, Preservica). DAMS that do offer PIDs (DSpace) tend to offer only fee-based PID options (Handle, DOI) rather than PIDs that are free (ARK, PURL, URN, local permalinks), slowing adoption, especially by institutions in the global South. The archival world is supported by a long tail of older, PID-unaware DAMS that continue to output unstable links, which create extra work and potentially disappointed users for aggregators such as Calisphere and DPLA.

Once objects are persistently findable, it is possible to achieve better connections between content and context. A downstream use of a digital object may not need to fully reproduce the entire context of a collection, but with some minimum information such as collection name and a persistent identifier leading back to the item in context, relationships can be better preserved.

Lean on widely-used standards, systems, and solutions.

Institution-specific solutions have value, both for specific needs and as proof of concept. However, they also promulgate two problems: unequal access to collections, and creating walled gardens that are

potentially less interoperable with other systems. This also makes success more difficult. But many of those difficulties could be solved by using common and domain-specific standards, which can help ensure that data is usable in the future, and, crucially, that subtle but fundamental conceptual requirements are being met.

In addition to standards, large-scale, even national, infrastructure and participation/membership structures that are accessible for institutions with related missions can help spread understanding and adoption. DACS, for example, provides excellent guidance for understanding and thinking about the role of context. However, the standard may not be as recognized outside of the archival community because this is not a core disciplinary concern. Sharing knowledge and infrastructure and being conversant in the disciplinary and standards conversation can often help software developers model and promote users' requirements.

What success can look like

To move from principles to solutions, what might success look like? Although we believe that system-agnostic principles are the most important step at present, it's helpful to have concrete examples to make principles easier to understand. This section provides examples of projects that implement some of the principles and a set of recommendations for system designers and developers.

One example that leans in the right direction to fulfill these principles, and that has a distinct advantage of simplicity and scalability, is the prototype co-developed by the [Orbis Cascade Alliance](#) and the University of Virginia's [Institute for Advanced Technology in the Humanities](#) (IATH) through a 2011-2014 National Leadership Grant from the Institute for Museum and Library Services (IMLS).³ This solution relates [Archives West](#), a database of over 30,000 EAD finding aids from Washington, Oregon, Idaho, Montana, and Utah, with the digital objects harvested for its DPLA hub. The digital objects appear as a group in the finding aid at the collection level, and a separate page retains that collection level information:

³ <https://www.imls.gov/grants/awarded/lg-07-11-0290-11-0>

Albert Henry Barnes photographs, circa 1895-1920 PDF XML

Overview of the Collection ▾

Photographer Barnes, A. H. (Albert Henry), 1876-1920

Title Albert Henry Barnes photographs

Dates circa 1895-1920 (inclusive)

Quantity 2000+ glass plate negatives
Nitrate negatives

Collection Num... PH0542

Summary Photographs of Western Washington including scenic views and images of early homesteaders

Repository [University of Washington Libraries, Special Collections](#)
Special Collections
University of Washington Libraries
Box 352900
Seattle, WA
98195-2900
Telephone: 206-543-1929
Fax: 206-543-1931
speccoll@uw.edu

Access Restricti... The collection is open to the public.
Nitrate and glass plate negatives are not available for viewing.

Additional Refe...

Languages English

Digital Objects



302 total - see all

Figure 4: Albert Henry Barnes photographs finding aid with associated digital objects.
(<https://archiveswest.orbiscascade.org/ark:/80444/xv76599/>, accessed 2021 June 11)

ARCHIVES WEST ORBIS CASCADE ALLIANCE

ABOUT CONTACT HELP

Overview of Collection

Albert Henry Barnes photographs

Creator: Barnes, A. H. (Albert Henry), 1876-1920

Dates: circa 1895-1920

Quantity: 2000+ glass plate negatives Nitrate negatives

Collection Num... PH0542

Summary: Photographs of Western Washington including scenic views and images of early homesteaders

Repository: [University of Washington, Special Collections](#)

Languages: Collection materials are in English.

Associated Cultural Heritage Objects

Records 1 to 24 of 302

1 2 3 4 ... Next 13



West entrance to Wright Park, Tacoma, Washington, ...

Point Defiance Park, Tacoma, Washington, ca. 1910.

Southwest slope of Mount Rainier showing Gilbralte...

Station house, Point Defiance Park, Tacoma, Washin...

Figure 5: Albert Henry Barnes photographs finding aid with all associated digital objects.
(<http://archiveswest.orbiscascade.org/do.aspx?ark=ark:/80444/xv76599>, accessed 2021 June 11)

A digital object from the same collection appears in the University of Washington's DAMS with a link to the corresponding finding aid's Archival Resource Key (ARK):

UNIVERSITY LIBRARIES
UNIVERSITY OF WASHINGTON

[Digital Collections](#)
[Browse](#)
[Special Collections](#)
[Exhibits & Guides](#)
[Ordering & Use](#)
[★ Saved Items](#)

[Home](#) > [Barnes \(Albert H.\) Photographs of Western Washington, 1895-1920](#) > [West entrance to Wright Park, Tacoma, Washington, approximately 1910.](#)
[West entrance to Wright Park, Tacoma, Washington, approximately 1910.](#)

Item Description

Title	West entrance to Wright Park, Tacoma, Washington, approximately 1910.
Photographer	Barnes, A. H. (Albert Henry), b. 1876
Date	between 1905 and 1915
Notes	Shows park grounds with two classical statues of women.
	On sleeve of negative: Tacoma, Wright Park, West entrance.
Subjects (LCTGM)	Parks--Washington (State)--Tacoma Public sculpture--Washington (State)--Tacoma
Subjects (LCSH)	Wright Park (Tacoma, Wash.)
Location Depicted	United States--Washington (State)--Tacoma
Digital Collection	Albert Henry Barnes Photographs
Order Number	BAR261
Ordering Information	To order a reproduction, inquire about permissions, or for information about prices see: http://www.lib.washington.edu/specialcollections/services/reproduction/reproduction Please cite the Order Number
Citation Information	University of Washington Libraries, Special Collections, [Order Number or Negative Number]
Negative Number	Barnes 61
Repository	University of Washington Libraries, Special Collections
Repository Collection	Albert Henry Barnes Photograph Collection, PH Coll 542
Repository Collection Guide	http://archiveswest.orbiscascade.org/ark:/80444/xv76599
Object Type	Photograph image
Physical Description	Silver gelatin, b/w
IIIF Manifest	https://digitalcollections.lib.washington.edu/iiif/info/barnes/0/manifest.json
IIIF Image	https://digitalcollections.lib.washington.edu/digital/iiif/barnes/0/full/full/default.jpg

Figure 6: Item from Albert Henry Barnes photographs in University of Washington DAMS, showing ARK link to finding aid. (<https://digitalcollections.lib.washington.edu/digital/collection/barnes/id/0>, accessed 2021 June 11)

This solution is based on the inclusion of the finding aid ARK in every digital object record and an OAI set harvest link in the finding aid.⁴ It meets the criteria of being scalable, simple to implement, and relies on an easy to understand data model. The information travels with items contributed from the Orbis Cascade harvester to DPLA. However, it does not provide the degree of context to individual digital objects described in our final criteria.

⁴ For more details, see the [Orbis Cascade EAD Best Practices](#) and [Dublin Core Best Practices](#). The documentation and code for the harvester that enables this solution is available at Github. Sam Meister's presentation on the project, which includes both additional details and screenshots from a beta product, is available at <https://www.slideshare.net/samalanmeister/the-crosssearch-and-context-utility-contextualizing-digital-content-and-associated-encoded-archival-description-finding-aid-metadata-in-the-northwest>

A less simple approach, but one that engages more successfully the notions of inheritance of description within a collection, is Artefactual Systems' AToM (Access to Memory). Originally developed with resources from the International Council on Archives, the data model for AToM is based on the findings of the Expert Group on Archival Description (EGAD) (ICA, 2012).

A digital object from within a collection of family papers appears as a standalone object:

Maki family: Oscar and Ju...



Clipboard

[Add](#)

Explore

[Reports](#)

[Browse as list](#)

[Browse digital objects](#)

Export

[Dublin Core 1.1 XML](#)

[EAD 2002 XML](#)

Finding aid

[Download](#)

Related places

[Canada » Ontario » Sudbury, Greater](#)

Physical storage

Shelf: DD-10/11/12/13

Identity area

Reference code	ON00120 016-.1-1-2-6
Title	Maki family: Oscar and Julia Koivula (sister of Anna). Lilja sitting on lap. Oiva and Oscar (sons) at back. Missing: Tiimi (daughter) Matti Kantokoski with wife Anna (Koivula) and children Eeva Annikki and Veikko Vesa
Date(s)	• [ca. 1929] (Creation)
Level of description	Item
Extent and medium	1 photograph : b&w copy print ; 15 x 10 cm mounted on paper 22 x 27 cm

Context area

Name of creator	Chisholm, Ann Eva (nee Kantokoski/Koski) (1924 - 2000)
Repository	City of Greater Sudbury Archives

Content and structure area

Scope and content	One annotated print copy of a photograph of the Maki and Kantokoski families. In the front row from left to right are: Oscar Maki, Julia Maki (holding baby Lilja Maki), Eeva Annikki Kantokoski, Anna Kantokoski, and Veikko Vesa Kantokoski. In the back row ... »
-------------------	--

Conditions of access and use area

Finding aid	kantokoski-koski-koivula-korpela-family-3.pdf
-------------	---

Allied materials area

Existence and location of copies	Item has been digitized. Prints of image may be requested. Email archives@greatersudbury.ca for more information.
----------------------------------	---

Access points

Place access points	• Canada » Ontario » Sudbury, Greater
---------------------	---

Digital object metadata

Media type	Image
Mime-type	image/jpeg

Figure 7: Image of Maki family with item-level metadata

The digital object display also shows the hierarchy of items, folders, sub-fonds, and fonds:



Figure 8: Image of Maki family with metadata hierarchy

By selecting within that hierarchy, the inherited collection context is evident:

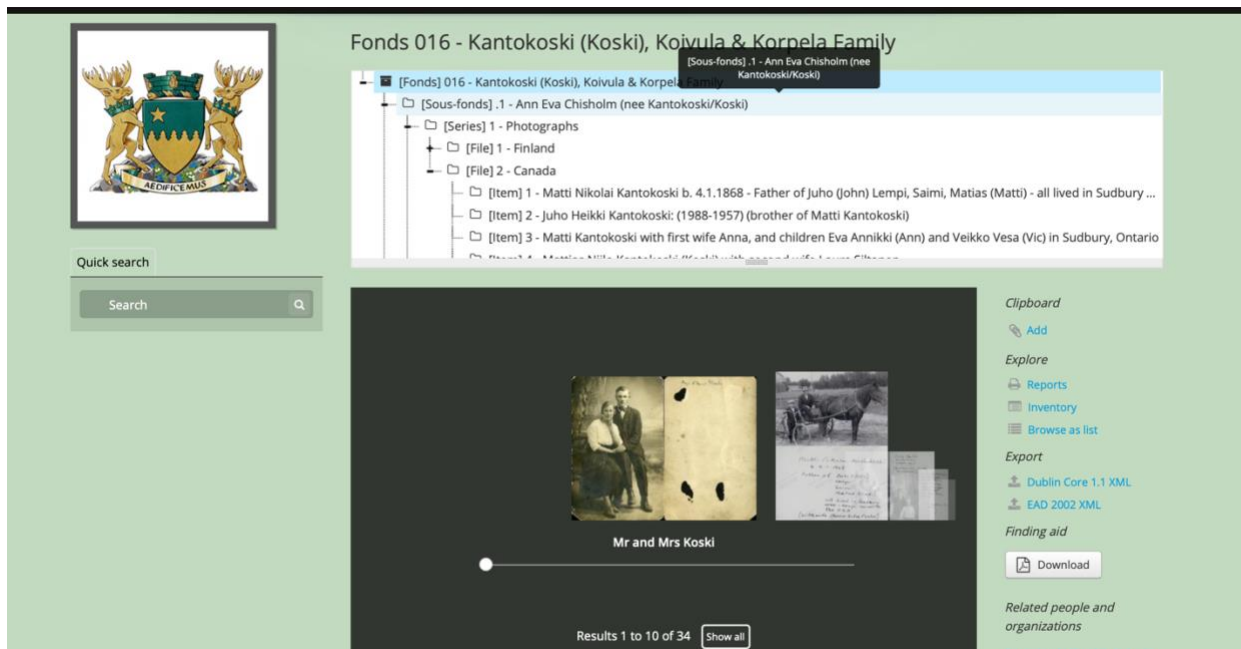


Figure 9: Metadata hierarchy for Kantokoski (Koski), Koivula & Korpela Family fonds

And the full collection-level description is available:

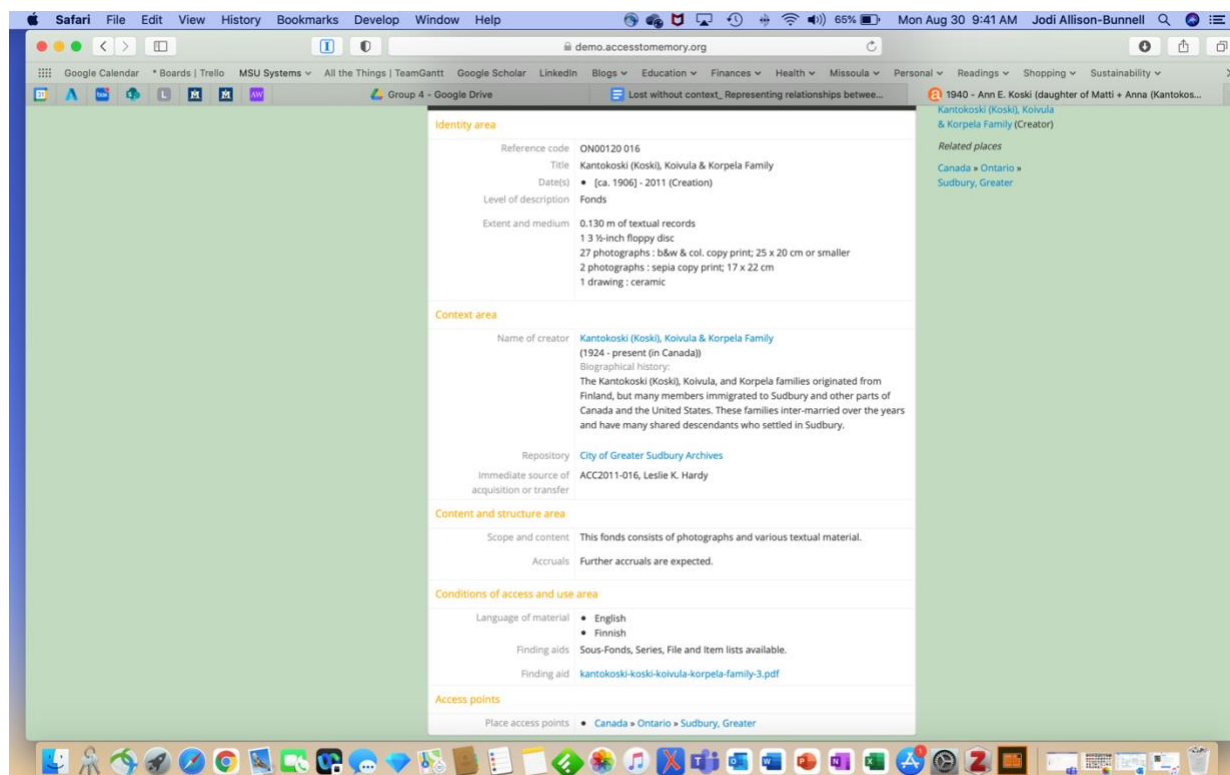


Figure 10: Collection-level metadata for Kantokoski (Koski), Koivula & Korpela Family fonds (<https://demo.accesstomemory.org/maki-family-oscar-and-julia-koivula-sister-of-anna-lilja-sitting-on-lap-oiva-and-oscar-sons-at-back-missing-tiami-daughter-matti-kantokoski-with-wife-anna-koivula-and-children-eeva-annikki-and-veikko-vesa>, accessed 2021 August 30)

AToM successfully engages inheritance of collection- and sub-collection-level descriptions. It adheres very well with many of the principles that we have outlined above. Whether the current presentation of the digital object and the collection context is understandable to end users is another question that may require further investigation. And while AToM is scalable in Canada (where it is used for the provincial archives, among others), it also requires meeting a high standard of standards compliance (to the Rules for Archival Description, or RAD) that may be a very high bar for institutions in the United States and thus less scalable.

What is needed for the future

These examples are successful because they follow the principles outlined above. They value context, make relationships clear, use aggregate description, and are based on the importance of maintaining historical integrity. They also rely on existing standards and often on widely-used systems so that materials can be shared and systems can be more easily migrated in the future.

But even within these examples, archivists and technologists had to go through a difficult process of translating principles into systems. The existing encoding standard, EAD, and the content standard, DACS, do not provide a clear formula for how to model archival data and its dependence on notions of

inheritance in a way that could be immediately actionable by a software developer or that can be implemented in many existing DAMS. In the future, as our users' expectations for archival materials online continue to grow, the archival profession would be well-served by a data model for archival representation that makes it possible to create systems for archival contents and contexts. For this archivists need to come to a firmer consensus on what inheritance means in archival description and work to better center archival standards around this idea.

Conclusion

For far too long, digital versions of items in archival collections have suffered upon being imported into DAMS, and from there into downstream object systems. The importing systems have no data models to account for archival context and persistent object links, resulting in an inevitable loss of richness, context, stability, and user experience.

The full complexity of the challenges of digitizing archival collections has in the past been more the subject of admiration than of determined, focused, and prioritized problem-solving. It does not need to be that way. DAMS designers and archivists can work together to create a more consistent and richer digital experience through modifications of data models and software to incorporate a small number of crucial archival context elements. As a result of conversations and work among archivists and others, at long last a few systems are beginning to show signs of accommodating archival context and persistent links. Making more DAMS consistent with archival representation, especially those used at scale and at resource-poor institutions, could have a large impact. The most critical next step? Having system designers (and archivists themselves) better understand the affordances of archival description. This will allow us to move forward with solutions that enhance discovery and access for all users and all cultural heritage materials.

References

- DeRidder, Jody, Amanda Presnell, and Kevin Walker. 2012. "Leveraging Encoded Archival Description for Access to Digital Content: A Cost and Usability Analysis." *The American Archivist* 75 (1): 143–70. <https://doi.org/10.17723/aarc.75.1.5641v61p422u0u90>.
- DPLA Archival Description Working Group. 2016. "Aggregating and Representing Collections in the Digital Public Library of America." <http://bit.ly/dplaCollections>.
- Duff, Wendy M., Emily Monks-Leeson, and Alan Galey. 2012. "Contexts Built and Found: A Pilot Study on the Process of Archival Meaning-Making." *Archival Science* 12 (1): 69–92.
- Ellis, Shaun, and Maureen Callahan. 2012. "Prototyping as a Process for Improved User Experience with Library and Archives Websites." *The Code4Lib Journal* no. 18 (October 3). <https://journal.code4lib.org/articles/7394>.
- Force, Donald C., and Randy Smith. 2021. "Context Lost: Digital Surrogates, Their Physical Counterparts, and the Metadata that Is Keeping Them Apart." *The American Archivist* 84 (1): 91–118.

- Greene, Mark, and Dennis Meissner. 2005. "More Product, Less Process: Revamping Traditional Archival Processing." *The American Archivist* 68 (2): 208-263.
- Hedstrom, Margaret. 2002. "Archives, Memory, and Interfaces with the Past," *Archival Science* 2: 21-43.
- International Council on Archives, Expert Group on Archival Description. "ICA Entity Model." <https://www.ica-atom.org/doc/UM-1.3>, accessed 2021 August 30.
- Lee, Christopher, A. 2011. "A Framework for Contextual Information in Digital Collections." *Journal of Documentation* 67 (1): 95-143.
- Matusiak, Krystyna K. 2017. "User Navigation in Large-Scale Distributed Digital Libraries: The Case of The Digital Public Library of America." *Journal of Web Librarianship* 11 (3-4): 157-171.
- Miller, Larisa. 2013. "All Text Considered: A Perspective on Mass Digitizing and Archival Processing." *The American Archivist* 76 (2): 521-541.
- Nesmith, Tom. 2005. "Reopening Archives: Bringing New Contextualities into Archival Theory and Practice," *Archivaria* 60 (Fall): 259-274.
- Society of American Archivists Technical Subcommittee on Describing Archives: A Content Standard (SAA TS-DACS). 2021. "Requirements for Multilevel Descriptions." *Describing Archives: A Content Standard*. Version 2021.0.0.2. https://saa-ts-dacs.github.io/dacs/06_part_1/02_chapter_01.html#requirements-for-multilevel-descriptions.
- Trace, Ciaran B. 2020. "Maintaining Records in Context? Disrupting the Theory and Practice of Archival Classification and Arrangement." *The American Archivist* 83 (2): 322-372.
- Yakel, Elizabeth. 2003. "Archival Representation," *Archival Science* 3 (1): 1-25.
- Zhang, Jane. 2012. "Original Order in Digital Archives." *Archivaria* 74: 167-193.

Maximizing Good: An Inquiry-Based Approach to Minimal Description for Online Archives

Sarah Dorpinghaus, Cory Lampert, Rebecca Pattillo, and Kyna Herzinger

Abstract: Minimal descriptive practices have been embraced by archives over the past fifteen years for their efficiency and practicality. This paper explores the strengths and weaknesses of minimal description within the context of digitized collections and evaluates them against the assumptions made by cultural heritage professionals. It considers whether minimal description provides digitized collections with sufficient metadata to meet MPLP’s user-centered goals of improving access, sufficient context to ensure collections are understandable within their digital environments, and sufficient framework to facilitate data exchange across systems, all while considering MPLP within archival ecosystems that impact labor and resource allocation. The authors offer a set of questions under four themes that challenge these assumptions and promote critical evaluation of professional norms related to minimal description of digitized collections. Recommendations are presented that realign methods to develop nuanced strategies that maximize our ability to steward our collections, respect our labor, and serve our users.

Introduction

“As you dig your teeth into your assumptions, your teeth become sharper. You can dig deeper. It’s not easy, but it is worth it. The truth, as they say, hurts. But they also say it sets you free.”

— Vironika Tugaleva, *“The Art of Talking to Yourself”*

As archivists, librarians, and digital repository managers, we wear many hats. We work to preserve and protect collections, we work to connect users to the knowledge contained in those collections, and we work to transform collections through technology as we anticipate future research methods. These key—at times aspirational—functions energize our duties even as they compete for time and resources (as they so often do!). This makes critical evaluation of our own methods difficult. Yet to succeed in our multifaceted roles, we must be willing to interrogate our assumptions and to build evaluation into our most basic processes. For that reason, our primary goal is to foster a culture of inquiry and exploration, even if that means asking questions that may be complicated to answer or questions that may not have any answer at all.

One of our profession’s prevailing assumptions is that minimal description is a sensible—if not preferable—practice with the gains in efficiency well worth its immediate shortcomings in intellectual control or user access. This area is ripe for critical reflection as little consideration has been given to the

effect that minimally processed collections may have on subsequent curatorial activities like digitization and online representation. As we interrogate our daily practices and challenge our assumptions about minimal description, which is best articulated in the methodology of “More Product, Less Process,” we are confronted by the messy realities of our work. Indeed, we gain a deeper understanding of the strengths and weaknesses of minimal methods with an eye toward developing more sophisticated approaches. Rather than relegate descriptive practices to one of only two choices—minimal or full—we can develop nuanced strategies that maximize our ability to responsibly steward our collections, respect our labor, and serve our users as partners in their research endeavors.

Since its arrival over fifteen years ago, More Product Less Process (MPLP) has informed the way that archivists navigate their work, making its way into the professional vocabulary and leaving an indelible mark on the literature. Indeed, few archivists are unaware of Mark Greene and Dennis Meissner’s “Low-Calorie, High-Fiber” approach, which sought to minimize backlog and introduce baseline metrics for processing. As they aptly described the challenges that many archivists faced, the duo made a case that minimal practices ought to be the go-to model for all arrangement, description, and preservation activities unless otherwise warranted. In so doing, they prioritized access to collections over unsustainable professional practice, ultimately shifting archival focus to user needs. Quite simply, they challenged their profession with a compelling question: what is the least we can do to get the job done in a way that adequately meets user needs both now and in the future (Greene and Meissner 2005)?

Even though its reception in formal venues has been generally positive, MPLP remains debated in casual discourse and even the occasional peer-reviewed publication (Cuervo and Harbeson 2011; Phillips 2015; Van Ness 2010). Minimal descriptive practices have appealed overwhelmingly to workplace sensibilities that venerate efficiencies, but it has also triggered thoughtful consideration of what has been lost. Most notable is Cox’s discussion of “maximal processing,” which considers the potential long-term impact of minimal description on discovery (Cox 2010). These efficiencies have prompted minimal practices to expand and evolve so that what was once proposed as a processing methodology for modern, paper-based collections has since morphed into a toolbox of practices that archivists have connected to nearly every aspect of archival administration and, most recently, digitized collections (DeRidder, Presnell, and Walker 2012; Evans 2007; Jackson 2012; Miller 2013; Sutton 2012).

This paper explores the strengths and weaknesses of minimal descriptive methods within the context of digitized collections and evaluates them against the assumptions made by cultural heritage professionals. It considers whether minimal description provides digitized collections with sufficient description to meet MPLP’s user-centered goals of improving access, sufficient context to ensure collections are understandable within their digital environments, and sufficient framework to facilitate data exchange across systems, all while considering MPLP within archival ecosystems that impact labor and resource allocation.

Impact of Minimal Description on Digitized Collections

At the same time archives have embraced the MPLP framework, funding agencies have prioritized online access to collections as researcher expectations for the same have grown. Administrators and practitioners have responded to these pressures, at times with little consideration of the fundamental assumptions that have informed decisions about the tools and systems or workflows and practices. This, in turn, has led to several problems: a glut of digitized resources with sparse descriptions, online

interfaces that place users into a confusing world of hierarchical description, rigid systems and workflows that are unable to adapt to changing technologies and user needs, and workers who continue to generate scans with minimal description but without understanding how or if their efforts are useful to researchers. The four statements below identify specific assumptions that often inform decisions about online access to archival collections and explore the unintended consequences when put in practice.

Assumption #1: Digitization of minimally processed or unprocessed collections allows for quick online access. Increased online access is equivalent to increased (meaningful) use.

Although digitization unequivocally gives users and researchers greater access to materials that would otherwise require an in-person visit, how effective is online access when discovery is limited due to minimal description? Case studies and organizations have made the argument that any online access is better than no access, an insight shared broadly across the profession during the recent pandemic (Dorpinghaus et al. 2019). Yet this is often coupled with the attitude that digitization is a magic bullet for access as cultural heritage workers abdicating further responsibility and additional stewardship once it is digitized and online. Those responsible for stewarding these collections conceptualize digitization as a means to an end rather than foundational jumping off point for enhanced access.

Digital records with detailed, item-level metadata receive more use due to better search retrieval. Whereas complex digital objects, particularly those with minimal description, lack that same level of access. As large-scale and full collection digitization projects have become the norm, adoption of practices that recycle minimally processed archival collection metadata and represent digital objects based on finding aid structures (one folder = one digital object) the item-level description is replaced with metadata that may never have been intended for the digital environment. Context assumed in one metadata standard is often missing in another. Users may only retrieve an object's parent record which is necessarily described less specifically than items within the folder. They are then left with the task of navigating a long list of "child objects" sometimes with no more description than a date range or vague title. While some users may expect this as "all part of the archival research experience" when working in-person with physical collections, little data has been collected about online users' tolerance for extensive clicking and review of digital images. What we do know is that users often use what they can easily find.

In addition to ease of access, one of the prevailing benefits of researching with digitized online content is the perceived ease of search and discovery. While there may be data about a particular place, event, or person in a collection of, for example, letters or diaries, minimal description means that these digital objects may not appear in search results despite containing information useful to the researcher. While the authors are not advocating that every diary entry receive a thorough summary, it is useful to consider how existing description will lead to successful search and discovery when selecting a collection for digitization for online access.

With this in mind, what research opportunities do users miss out on when discovery is difficult? "With minimal processing, we are creating a whole new generation of hidden history." (Cox 2010) What does it mean if many users either walk away frustrated and without resources or with something that is "close enough" when the actual gem remains hidden in the deluge of digitized content? How do our

decisions on processing and digitization impact whole fields of potential intellectual exploration? As underrepresented voices and community collections are prioritized for digitization, what does our level of metadata investment say about our commitment to responsible stewardship of these cultural resources? It is time to include a digitally focused metadata strategy in all decisions about processing and digitization. Avoiding these decisions “for now” and pushing this hard work “down the road” or “downstream in the workflow” assumes a vague future scenario when these answers will magically materialize. In the meantime, more and more digital objects are created every day.

Assumption #2: Users understand how to navigate archival discovery tools and description well enough to successfully conduct their research online.

Do users understand the nuances of archival description and context within the online environment? While a small percentage of researchers can answer in the affirmative, many users are not familiar with the concept of a finding aid or hierarchical description. Discovery of digitized archival resources is limited in multiple ways, but particularly when it comes to users’ level of digital proficiency and ability to navigate complex relationships between digital objects.

Online archival collections are used by a range of users with different needs, expectations, and research skills. From undergraduate students seeking primary resources for a course assignment to a community member researching their family’s history to a postdoctoral fellow doing in depth research on a specific topic, individuals are using online resources differently. Some are seeking to simply satisfy the requirements of an assignment, others are looking for the mention of a specific name or place, while others are looking to better understand the landscape of a topic during the context of a particular time and place. Each of these users come with a different set of understanding of how archives in an online environment function and are willing to put in different levels of work to find what they need.

In the pre-online period, users would visit the physical archive, assuming they could work around the barriers to researching in person, where an archivist could provide a level of scaffolding to help find relevant resources. However, as more collections are available online, the door is open to new users who arrive with little or no foundational skills in researching with archival material, to no fault of their own. This does not present a problem if the digital library interface is intuitive and metadata is sufficient for successful discovery. Knowing that is not always the case, the onus is placed on the researcher to actively seek out the archivist for help. Considering power dynamics and lack of time and other resources, users may not often contact the archives for assistance.

Other users have needs beyond online discovery and access. Research methodologies in the Humanities and STEM alike require access to large datasets that can be used for computational processing and other forms of analysis. In turn, archives are experiencing a rise in researchers seeking large dataset downloads or computational tools as part of online digital libraries (Green and Courtney 2015). Archives are often at a loss on how to meet these needs as mainstay tools and systems have yet to add such features.

To what extent are our systems (and the extent of our metadata) designed for the “super users” that are steeped in the world of archival research? Or are they actually designed for archivists and to work within our existing practices rather than provide the best experience for users? Is the scaffolding we provide appropriate and meeting users where they are (e.g., hover text rather than a five-minute video tutorial)?

A single digital library system cannot be built to meet the needs of every user. How can we adjust our systems, descriptions, and workflows to meet the varying needs of most users? Whose needs are prioritized when making decisions regarding description and digitization? To explore this problem further, we need to plan collaboratively and strategically, and seek feedback from users while preferencing their needs over the easiest or most cost-effective solution from the archive's perspective.

Assumption #3: Tools and systems will improve over time to solve known interface problems and increase interoperability.

Although users of online archival collections are diverse in experience and needs, they are united in a desire for a seamless, efficient, and simple research experience. Digital libraries have modeled themselves after online retail sites with consistent features like search filters, user accounts, and shopping carts. This reduces some of the entry barriers to using online archives. Yet, one key difference between online archives and online retail sites is that archives often must maintain some sort of relationship (either flat or hierarchical) between items. Some content management systems have done better at maintaining the hierarchical relationships of a collection, series, folder, and item (e.g., ArchivesSpace public user interface and ArcLight) while others (e.g., CONTENTdm and Omeka) work under the assumption that items will have sufficient metadata for discovery. In each of these environments, the role of minimal description greatly impacts successful discovery of archival resources that meet researcher needs. Or, just as importantly, allows researchers to quickly and accurately assess when a digital library does *not* have what they need. This in addition to the vernacular of archives ("finding aid", "scope and contents", "series") leaves some inexperienced or new users confused by digital libraries and hesitant to return.

Additionally, users must contend with navigating through the different tools and systems within a single archive. Users may start on a digital library and then be required to set up an account to request copies or perhaps navigate to a different tool to schedule an appointment for viewing in person. Each of these with a distinct look and feel. And this is just for access to archival resources; users may also be working with the library catalog, research guides, and have separate accounts for interlibrary loans. Likewise, if a researcher is utilizing digitized collections from different institutions, they are likely to encounter differences between those systems in regards to navigation, faceting, hierarchy, and levels of description.

How does this lack of consistency between digital library interfaces disadvantage users? What role does the home-grown or highly customized digital library play in a seamless research experience? How could user experience improve if minimal description is no longer acceptable for the online environment? It is time for the profession to explore breaking from the constraints of minimal description and traditional archival description structures for online digital content altogether.

Assumption #4: We can tackle that in a future phase. Or: Minimal description, while not ideal, is a necessary/adequate way to deal with persistent resource limitations.

When physical collections are in need of more detailed description, digitization can be seen as a salve to processing the physical collection. Particularly in visual resources collections, digitization is seen as

a way to gain intellectual control via item-level description of the collection, but this perpetuates disjointed decision-making often resulting in circular thinking. Greene and Meissner (2005) explain,

One of the first questions to ask in any digitization project is “Does the intellectual quality of the source material warrant the level of access made possible by digitizing?” One can posit this as a chicken/egg problem—how do we know if the collection is good enough to digitize if we haven’t already described it to the item level? But more practically, if arrangement and description of the analog material depend on an initial assessment of the value (or intellectual quality) of the collection in the first place, then finely processed collections will by definition be good candidates for digitization and require less additional descriptive work.

It is time to include a digitally focused metadata strategy in all decisions about processing and digitization and address the chicken and egg question head on earlier in the curation process, as suggested in the OCLC Report, “Total Cost of Stewardship” (Weber et al. 2021). This approach not only addresses prioritization needs, but overcomes cases where the MPLP or minimal description approach to digitization ends up as an excuse for under-resourced and understaffed institutions to continue to justify their lack of investment in archival labor. While new methods can be applied to newly acquired collections, previously processed collections may require more difficult decision-making about revisiting description levels and resource allocation. Fortunately, many information professionals have begun strongly advocating for the need to do reparative work on collections, including redescription of materials. This is an optimum time for a reflection on the past decade’s trend to prioritize digitization and online access above all else.

This is a call for change not only to cultural heritage institutions but also to those funding that work. The over reliance on grant funding for many aspects of our labor continues to create an unfair dependency on contingent and precarious positions, putting undue stress on information professionals (Rodriguez, et al. 2019). Further, reliance on piecemeal grant funding (especially for projects that utilize minimal description or MPLP practices) for specific projects acts to mollify the larger issues of under-valued and under-paid labor, few permanent positions, and lack of resources. Have we used MPLP as a short-term fix to long-term problems rather than devoting the time to evaluate and create ways to relieve the inequitable labor issues in our field, advocate for better funding, and make our labor more visible to our users and stakeholders (Williams 2016)? When digitization projects rely on time-based funding, often minimal description is the necessary route for completing the project within the scheduled parameters of the grant. Additionally, understaffed institutions may rely on minimal description workflows to complete projects, particularly when there may be one or a small handful of employees dedicated to digitization. As information professionals continue to find workarounds to a lack of investment in our repositories and our labor, we and our users are disadvantaged. Any large-scale digitization project that utilizes minimal description should be evaluated closely to ensure that the resources dedicated to the project are truly resulting in increased discovery and useability of the collection(s) while supporting ethical labor practice for the information professionals tasked to do the work.

Changing Practice

The challenges that have emerged from the assumptions identified above are widespread and have deep roots. As such, they demand creative solutions driven by tough questions and routine evaluation of existing practices. As we challenge individual, institutional, and professional norms, we can leverage ambitious practices that impact strategic development and planning, fair labor and ethics, technological tools, and assessment. We propose the following plans to accomplish this:

1. Create a plan for digitization at the point of curating and/or accessioning collections and have a hard conversation about prioritized work across the organization. Consider the life cycle of collections and align resource-heavy descriptive steps like processing and digitization together. Address descriptive needs in requests for external funding and eliminate digitization that does not include funding for description at the level needed.
2. Radically rethink the role of digital object representation (including finding aids) in the online environment. Work to retain valuable context, but also embrace the transformative nature of digital and online research by letting digital assets and metadata move out of, between, and through archival constructs. Where possible, implement systems that enable this transformation in both data models (linked data) and user interfaces (visualizing relationships and connections).
3. Do better for workers by thinking about the compounding effect of constant grant-funded and time-bound project deliverables. Avoid potential harm to contingent workers temporarily employed on digital projects and address workload for all employees, especially during (increasingly permanent) periods of resource scarcity.
4. In direct defiance of minimal practices, flip the whole system and invest in reparative description and work toward more inclusive metadata overall across systems. Develop metadata strategies that add value in broad and sustainable ways such as: alignment with the Santa Barbara Statement on Collections as Data (Padilla et al. 2019), normalization and cleanup of data for future interoperability and migration between systems, and/or assignment of RightsStatements.org controlled vocabularies.

Action & Evaluation

Many of the questions posed throughout this piece do not have clear overarching answers. Yet they do articulate the need for both reflection *and* action. There is work to be done. The following can be used to challenge assumptions and make decisions that push beyond ambitions to action and continuous growth.

Online Access

In order to appropriately utilize its limited resources of staff time and digital storage space, an archive must strategically select collections for digitization and online access. Ensure digitization is considered during donation, accessioning, and processing. Be realistic and honest when discussing digitization with donors, curators and collections managers, administrators, and users. Do not make promises without considering the short and long-term implications.

The foundational marker of research value will often be a motivator for digitization, but archives should also recognize reasons *not* to digitize or to postpone digitization and online access.

- Does the existing descriptive metadata provide enough context for understanding the items when in the online system?
- Does the existing descriptive metadata include natural language, controlled vocabularies, or full text transcripts in searchable fields that will allow for adequate discovery?
- If not, what is a realistic estimate of how much effort is required to expand the metadata sufficiently? Can aspects of this work be automated? Who will be responsible for this work and do they have the capacity? How will this impact other priorities?
- Has the collection been processed to a level that merits digitization? If there is little meaningful description do not digitize. Stop the cycle.

Usability and Navigation

Rather than prioritizing expedient online access, seek a balance between access, discovery, and positive user experience for all users. If we fail to do this, our digital libraries will become confusing networks of millions of digitized objects that may be accessible but remain largely undiscoverable for most users. Many archives are already standing on this precipice and find themselves, toes perched over the edge, realizing something must change.

Now is the time to pause and reflect, to gather data and consider long-term implications.

- What do we know about our users? Take time to gather feedback and conduct user experience studies, no matter how small in scope. Document findings and use them to inform and prioritize iterative interface changes that benefit the majority of users.
- What technical methods can be leveraged to extract useful information from the finding aid, such as collection notes and series and subseries information, for display at the item-level view to improve use of collections with minimal description?
- How can we collaborate with allied professions to update standards and design tools that offer a more cohesive, intuitive, and meaningful experience for a majority of users?

Tools and Systems

There is no one system that addresses all of the challenges in online discovery. The diversity of organizational needs has led to a plethora of digital library systems that prioritize different basic elements of archival description and access. It has been resource intensive to develop and customize these systems, many of which attempt to support minimal description, and it is even more intensive to transition out of these systems into something different. Rather than attempt to design tools and systems that support traditional archival practices, consider how discovery and meaningful use can be improved by breaking away from the rigid structures of conventional description.

- How can tools and systems prioritize features to support linked open data and the use of collections as data? How well is your data optimized for discovery outside of the local system? Can data be easily shared via an open API or SPARQL endpoint?
- Can we build systems that provide discovery and meaningful use of digitized archival materials without relying on the structure of the finding aid? Can the information from the collection guide be structured differently for online access?
- When the system is determined to be the problem, can an improvement in the data itself solve any of the issues? Would data clean-up help users more than adding a new interface? Could centralizing description and reusing data across functions improve discovery as much as a new interface? How ready is your data to move across systems if this is the only solution?
- When choosing a new architecture, many organizations think they are unique in their struggles and must build a unique system to fix the problems. Is this really true? What are the implications of selecting a homegrown development path: in the context of sustainability, interoperability, and getting locked into yet another “custom” way of doing things? Rather than invest in highly customized solutions, cross-institutional partnerships must be strengthened and community collaboration increased to build open source and interoperable tools, and invest in iterative improvements to metadata and interface design.

Prioritization and Labor Issues

The profession is in the midst of a labor crisis. Our stated priorities and needs have vastly outpaced our resources and evidence of this is seen in the nearly ubiquitous backlogs. These growing backlogs have rarely even begun to address born digital archives and online access. So what questions can we ask to make better decisions about priorities and responsible assignment of work?

- Question existing priorities. Ask, “why these activities first”? Determine if they are clearly aligned with strategic planning and values of the organization. How often are priorities reconsidered? Whose voices are heard in the process? How can these priorities remain agile in a rapidly changing world?
- Challenge assumptions about new collections coming in. Is the organization realistically capable of providing access to this collection in a way users expect? What metrics could be developed using tools such as the recent OCLC report, “Total Cost of Stewardship” (Weber et al. 2021)?
- When institutions do not have the resources, are grants or project-based funding the only option? Is it possible to break free of the external funding treadmill and make long-term investments in labor to support description? If not, is it worth the harm these positions perpetuate?
- If relying on term positions is the only way to provide access, what guidelines are in place to ensure equitable labor practices? Are these aligned with national conversations in groups working on improving contingent labor practices such as the *Collective Responsibility* National

Forum on Labor Practices for Grant-Funded Digital Positions white paper (Tillman and Rodriguez 2020)?

In their support of MPLP, Greene and Meissner had argued that “a sign of professional maturity would be for us to own up to the limitations we work under” and then adapt our methods to align with those limitations. After all, simply doing the same thing while expecting different results, they reminded us, is the very definition of insanity (Greene and Meissner 2005). As a profession, we need to embrace this challenge, by critically evaluating the utility of what we do and realistically aligning our methods to develop nuanced strategies that maximize our ability to steward our collections, respect our labor, and serve our users.

References

- Cox, Robert S. 2010. “Maximal Processing, or, Archivist on a Pale Horse.” *Journal of Archival Organization* 8 (2): 134-148. <https://doi.org/10.1080/15332748.2010.526086>.
- Cuervo, Adriana P., Eric Harbeson. 2011. “Not Just Sheet Music: Describing Print and Manuscript Music in Archives and Special Collections.” *Archival Issues* 33 (1): 41-55. <https://www.jstor.org/stable/23100472>.
- DeRidder, Jody L., Amanda Axley Presnell, and Kevin W. Walker. 2012. “Leveraging Encoded Archival Description for Access to Digital Content: A Cost and Usability Analysis.” *American Archivist* 75 (1): 143-170. <https://doi.org/10.17723/aarc.75.1.5641v61p422u0u90>.
- Dorpinghaus, Sarah (Session Chair), Rachel Gattermeyer, Cory Lampert, and Rebecca Pattillo. 2019. “Digital Double Bind?: Exploring the impact of More Product, Less Process (MPLP) on digital collections.” Conference panel, Digital Library Federation Forum, Tampa, FL, October 15. <https://osf.io/q9hcf/>.
- Evans, Max J. 2007. “Archives of the People, By the People, For the People.” *American Archivist* 70 (2): 387-400. <https://doi.org/10.17723/aarc.70.2.d157t6667g54536g>.
- Green, Harriett E., and Angela Courtney. 2015. “Beyond the Scanned Image: A Needs Assessment of Scholarly Users of Digital Collections.” *College & Research Libraries* 76 (5): 690-707. <https://doi.org/10.5860/crl.76.5.690>.
- Greene, Mark A., and Dennis Meissner. 2005. “More Product, Less Process: Revamping Traditional Archival Processing.” *American Archivist* 68 (2): 208-263. <https://doi.org/10.17723/aarc.68.2.c741823776k65863>.
- Jackson, Tracy M. 2012. “I Want To See It: A Usability Study of Digital Content Integrated into Finding Aids.” *Journal for the Society of North Carolina Archivists* 9 (2): 20-77. <https://www.ncarchivists.org/vol-9-no-2-spring-2012/>.
- Miller, Larisa. 2013. “All Text Considered: A Perspective on Mass Digitizing and Archival Processing.” *American Archivist* 76 (2): 521-541. <https://doi.org/10.17723/aarc.76.2.6q005254035w2076>.

- Padilla, Thomas, Laurie Allen, Hannah Frost, Sarah Potvin, Elizabeth Russey Roke, and Stewart Varner. 2018. *The Santa Barbara Statement on Collections as Data*. <https://doi.org/10.5281/zenodo.3066209>.
- Phillips, Jessica. 2015. "A Defense of Preservation in the Age of MPLP." *American Archivist* 78 (2): 470-487. <https://doi.org/10.17723/0360-9081.78.2.470>.
- Rodriguez, Sandy, Ruth Tillman, Amy Wickner, Stacie Williams, and Emily Drabinski. 2019. *Collective Responsibility: Seeking Equity for Contingent Labor in Libraries, Archives, and Museums*. Institute of Museum and Library Services. https://laborforum.diglib.org/wp-content/uploads/sites/26/2019/09/Collective_Responsibility_White_Paper.pdf.
- Sutton, Shan C. 2012. "Balancing Boutique-Level Quality and Large-Scale Production: The Impact of 'More Product, Less Process' on Digitization in Archives and Special Collections." *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 13 (1): 50-63. <https://doi.org/10.5860/rbm.13.1.369>.
- Tillman, Ruth Kitchin and Sandy Rodriguez. 2020. *Collective Equity!: A Handbook for Designing and Evaluating Grant-Funded Positions*. <https://doi.org/10.26207/6p4a-md61>.
- Van Ness, Carl. 2010. "Much Ado about Paper Clips: 'More Product, Less Process' and the Modern Manuscript Repository." *American Archivist* 73 (1): 129-145. <https://doi.org/10.17723/aarc.73.1.v17jn363512j545k>.
- Weber, Chela Scott, Martha O'Hara Conway, Nicholas Martin, Gioia Stevens, and Brigette Kamsler. 2021. *Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/zbh0-a044>.
- Williams, Stacie. 2016. *Implications of Archival Labor*. Medium. <https://medium.com/on-archivy/implications-of-archival-labor-b606d8d02014>.

Playing to our Strengths: Self-Assessment Criteria for Access and Discovery in Small Archives

Stefana Breitwieser, Amanda Demeter, Sophie Glidden-Lyon, Amanda Murray, Lori Myers-Steele, and Kate Philipson

Abstract: Archives are often confronted with challenges related to funding, staffing, technology, and high expectations from administrators and users. For small archives, these challenges can have an enormous impact on access and discovery of collections. This paper explores these challenges in detail, as well as the strengths specific to small archives. The authors conclude that by leveraging these strengths, small archives can reframe their understandings of successful access and discovery. The paper also provides a series of questions for archivists who wish to audit access and discovery practices at their institution and reevaluate what it means to be successful within the particular context of their small archive.

Introduction

Many archivists are familiar with a common set of challenges: finite, often grant-based funding; complexities of implementing and integrating technology and technological workflows; high staff turnover in the archival field; limited funding for ongoing archival training and education; on-demand access to digitized material and other significant expectations from researchers and administrators; and difficulties with institutional communication and internal advocacy. For those working at small archives these challenges are even more acute. For example, staff turnover can feel all the more painful on a team of one or two, and if budget conversations start with how to keep the lights on, funding new technologies or initiatives may seem like a stretch. Taken together, these challenges can be overwhelming, and may prevent archivists from making real progress toward improving access.

Convened as part of the Lighting the Way Working Meeting, a team of archivists met to discuss small archives (meaning those with either a solo archivist or “lone arranger,” or just a few team members) and how they enable discovery and delivery to their collections. Discovery and delivery is defined as “what people and systems do to support finding, accessing, and using material from archives and special collections. Systems include not just software, but also workflows, paper forms, standards, and more” (Matienzo et al. 2020, 3). (For the purposes of this paper, the broader term “access” will be used in place of delivery.) This team found that these challenges are common, interconnected, and sometimes existential to small archives. What would it mean to acknowledge and move beyond them? How can small archives be empowered to begin brainstorming creative strategies for discovery and access by embracing their unique strengths and readily available tools? This paper aims to answer these questions by defining the shared challenges of small archival institutions, and suggesting recommendations for making an honest and positive assessment of their abilities.

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Cite as: Stefana Breitwieser, Amanda Demeter, Sophie Glidden-Lyon, Amanda Murray, Lori Myers-Steele, and Kate Philipson, “Playing to our Strengths: Self-Assessment Criteria for Access and Discovery in Small Archives,” in *The Lighting the Way Handbook: Case Studies, Guidelines, and Emergent Futures for Archival Discovery and Delivery*, edited by Mark A. Matienzo and Dinah Handel, 85-96. Stanford, CA: Stanford University Libraries, October 2021. <https://doi.org/10.25740/gg453cv6438>.

It is important to acknowledge the reality of these challenges and that archival work continues in spite of them. Working in a small archive comes with its own unique strengths and benefits including, but not limited to: smaller administrative overhead; closer connections with patrons and donors; the ability to redefine access by catering to a specific community; greater flexibility to experiment, make changes, and adapt processes; and more hands-on experience with the collections as a whole.

Smaller institutions should be encouraged to focus on what's within their reach, and to visualize what they have the capacity to do now, without waiting around for better funding or staffing that may never materialize. It is hoped that there can be a shift in what is expected of small archives within the culture of the profession and a recognition that what is best for large, well-resourced institutions is not necessarily what is best for small archives. Archivists should feel empowered to advocate for best practices that factor in the context of their institution, including specific needs, priorities, and resources. The following discussion explores these challenges and strengths in greater detail, and concludes with a series of questions to consider, for any archivist hoping to reframe some of the so-called shortcomings of working in a small archive into avenues for potential action.

Shared Challenges

With the expansive growth of utilizing online systems for discovery and access, collections will become increasingly invisible if they are not accessible online in some form, whether through descriptive records or digitized materials. With the growing expectation that collections are accessible online, smaller institutions struggle to find innovative ways in which to increase opportunities for discovery and access for their patrons. With their time and resources already stretched to handle every facet of archival work, small archives have to prioritize and allot limited resources on a daily basis.

Archives staff in small organizations are accustomed to needing a broad knowledge base and a flexible approach to the many facets of archival work. This work often includes the need to balance time between technical “back-end” tasks to make collection materials accessible, as well as working directly with those who use the materials. The expectation that the archivist has experience with every aspect of archival work and is able to educate staff members and patrons in running or using an archive requires the archivist to teach themselves new skills and engage in professional development opportunities. Identifying opportunities that speak specifically to the unique needs of small archives can be challenging in itself; even when supported by administration, a lack of funding and time limits both participation in development opportunities and implementation of what was learned. Scarcity of resources is again a culprit in preventing archivists from making collections accessible. In their article reporting on a broad survey of archival work, Buchanan, Gruning, Gursoy, and Barker (2017) identify a wide range of interactive tasks (primarily people-centered) and materials management tasks (object-centered) that archivists engage in as part of their regular roles (276-278). The survey of archivists identified resource scarcity (in funds, staff, and time) and communication and advocacy for archivists' work as their biggest challenges (278-280). These challenges were further corroborated through verbal anecdotes shared by the authors of this paper.

Additionally, turnover among staff, particularly when institutional knowledge is lost with an exiting staff member, is a substantial challenge, especially in terms of trying to create a system of discovery and delivery. Small archives can see high turnover rates, often due to a large number of term or part-time positions, and a lack of full-time and permanent positions. This increased staff turnover can lead to

inconsistency in the vision or goals for a small repository, cause difficulty in advocating for the value of the archive, and can substantially disrupt the services and systems utilized for access of collections. A lack of documentation can also easily lead to inconsistencies in accessioning, processing, and description work, making both internal and external access harder.

A perennial problem for almost all organizations is funding and budget constraints. In creating access, a lack of funding inevitably limits an organization's ability to sustain and support needed staff, as well as procure, implement, and maintain discovery and access systems (and the infrastructure necessary to support those systems). While administrators making decisions about budgets are often supportive in general of necessary archival work, budget decisions don't often reflect that support in terms of sustained funding streams. Flashy updates and systems are often the more appealing projects to fund than those that are essential to the basic documentation and maintenance needs of the archive. The internal priorities of maintaining an archive are often overlooked by financial decision-makers outside of the archive.

Forced to cover budget gaps and meet basic archival needs, archivists often look to external grant funding. While grant funding for specific projects or one-off programs is greatly helpful, applying for and accepting these funds often comes with its own set of challenges and expectations. Grant funding opportunities often closely follow current trends in the field that small organizations are not prepared to (nor necessarily have a need to) meet. Additionally, grant funding often requires investment of scarce resources or matching funds that a small organization will not be able to sustain, especially while other more basic needs have not been met. Crafting a successful grant application requires an immense time investment by archival staff, and often comes with attached requirements and expected outcomes. As intended, grant monies are finite but leave the organization to navigate how to sustain the requirements and outcomes in perpetuity. Expected outcomes and requirements, which can range from providing digital preservation measures, to creating extensive catalog records and metadata, to providing ongoing technical support and platform hosting, strain small archive staff and are unrealistic for their organizations. Furthermore, they may cause future scarce resources to be allocated in ways that do not reflect the basic needs of the repository. One-off programs and grants, such as those offering funding, staff, or services to implement exciting digitization and cataloging projects for increased access to collections, allow institutions to provide better discovery and access to specific collection materials; however, they do not necessarily aid the institution in maintaining a sustainable system for all their collections.

While making some collections more accessible, grant funding often places unsustainable strains on available staff and resources. Those same workflows and expectations cannot be replicated during the processing and cataloging of additional or future collections without a similar influx of funding. These funded projects tend to highlight and increase accessibility to some materials, while increasing the expectations of archival patrons that the same will be done for other collections, when there is no sustainable system for doing so. Managing a small archival program is no easy task and increasing expectations such as these does not make it any easier. Even if an institution can secure temporary or base funding and support for creating systems of discovery and access, additional persistent and long-term challenges exist and must be faced to achieve a sustainable system that works to make all collections discoverable and accessible to an organization's patrons. While all institutions may envision a future state of providing an interoperable system of discovery and access that both meets the needs

of users and is maintainable by staff, the approach to achieving that vision will most likely differ for small archival institutions.

Adopting archival technologies (such as archival management systems, digital preservation systems, online exhibit platforms, etc.) are a significant challenge to small institutions. New technologies are enticing, and administrators, in particular, may see technocratic solutions as the key to addressing other issues facing small archives. Certainly, to an extent, this is true. For instance, basic scripting can reduce the burden of redundant, tedious workflows; databases allow for better and increased collections management; and online systems of outreach are a key access point for many institutions. For small repositories, however, scripting, as well as the creation and maintenance of collection management systems, databases, and online systems, require significant budgets and expertise — an expertise that is often missing from archival education, and a demanding request of an archivist who is already spread thin. Consequently, a lack of administrative and patron understanding regarding the extensive funding, time, and expertise needed for digital preservation and the creation and maintenance of online systems for access can create an environment wherein the archivist becomes overwhelmed by unrealistic expectations. If digital objects and online systems are popularly conceived of as ephemeral and instant, then patrons, donors, and administrators will have expectations for immediate discovery and accessibility of collections with no conception of the time, expertise, and resources required to digitize materials and provide online systems of access.

It would appear that communication and advocacy with administrators, donors, and patrons is key to addressing high expectations to the online accessibility of collections; however, communication and advocacy are particular challenges for small archives. Smaller archives often find themselves in institutions and administrative structures that do not understand the mission, scope, and impact of the archive. Siloed departments, expectations of digital access, misunderstandings of what archival work entails, and persistent public misconceptions about archival holdings and processes all contribute to this. Without a sense of what archival labor looks like, it is understandably difficult for users, donors, and administrators to know what a reasonable expectation for discovery and delivery is in a digital age when information appears to be effortlessly available at our fingertips. This requires the archivist to take on a further expanded role, that of the educator. Archivists of the small repository are often burdened with articulating to administrators, donors, and patrons the value of not only the archive, but of the archivist's job itself. As stated above, Buchanan et al. (2017) found communication to be the second most cited challenge for archival workers (278-280). One solo archivist who completed the survey shared that they were the first archivist at their institution, which meant that they spent much of their time "doing outreach and educating the archives' stakeholders about the purpose and benefits of the archives" (280).

Being able to effectively communicate with internal and external stakeholders and to advocate for an archive is closely connected to the ability to take part in decisions that are essential to the archive's operations. When archivists are not included in broader institutional conversations, they miss important opportunities to demonstrate value, build partnerships, and even build the collection. Therefore, small shops and their archivists also need to be able to communicate their needs and limitations. This can feel risky, but when administrators and other high-level stakeholders are unfamiliar with what is required for a successful, sustainable archival program, they may ask for features or projects that are beyond the scope of technological, staffing, and/or funding capacities. Being willing to articulate the needs of the archive and the community that depends on it is an integral

part of educating others, and advocating for your collections. An open and honest conversation can build good will.

Unique Strengths and Opportunities

Small archives operate on a scale that provides them with fundamental strengths and opportunities. This smaller scale is the source of the beneficial characteristics that will be explored here when considering archival discovery and access systems: a greater flexibility and adaptability, the opportunity to experiment more easily, the ability to apply new tools or technologies across collections quickly, and a more direct line of communication with administration, donors, and patrons. Because of their small scale, archives that are often under-resourced can instead utilize these strengths to their advantage in providing discovery and access services.

Small archives are more likely to have more horizontal, less hierarchical administrative structures, which can increase their flexibility and communication when it comes to institutional decision-making about necessary systems of discovery and delivery. While this is in tension with the challenge of communication and advocacy that was noted previously, archivists may find that there are a few distinct benefits that come with working in this type of administrative structure. Fewer layers of administration means that any required reviews and approvals can happen more quickly and directly. For example, a small archive obtaining approval for a policy document is unlikely to have to get feedback from multiple committees, departments, or administrators prior to implementation. More direct communication with fewer channels throughout an organization can also allow for an ease of experimentation with new access tools and methods. Additionally, less bureaucratic hierarchy within a small organization may help create a more collaborative approach to project approvals and advocacy. Where there is limited staff, archivists may have a louder, more unified voice in advocating for their needs around creating and providing access. Indeed, many small archives, particularly within historic houses and small museums, may find that some decision-making begins and ends with the archivist. Alternatively, when a small archive is housed within a larger institution, the archivist may be only one of a few voices in collection management, so administrators must rely on their specialized expertise. The archivist holds a position of trust – while administrators may not be able to fully support every initiative, the archivist has a direct voice in advocating for resources and guiding projects and programming. Just as importantly, this trust also allows the archivist to advocate against other projects that administrators prefer, but which do not align with the true needs of the archive or the realities of collection management. For example, at one small archive, an administrator was interested in optical character recognition (OCR) for handwriting powered by artificial intelligence. The archivist encouraged the administrator to wait until this technology was further developed, but used the opportunity to successfully advocate for additional technological infrastructure to support online access to digitized materials, arguing that this was a necessary precursor to potential future projects. The archivist provides an authoritative, consistent voice, drawing from direct, hands-on experience with the collection and its researchers.

Another strength of smaller archives that can be leveraged in creating a system of discovery and access is the ability of their archivists to form close relationships with, and better appreciate the needs of, donors, patrons, and the broader community which the collections represent and/or serve. In small archives, archivists may work closely with donors of collections; serve as the processors, describers, and digitizers of collections; and provide services to users, including answering reference emails,

scheduling appointments, helping people navigate catalogs and websites, assisting internal users, and more. These archivists are more likely to know exactly who their patrons are and can use that intimacy to develop a clearer sense of what those users want and need from the archive. As Colleen McFarland (2007) writes, "we must cultivate skills that help us 'sell' ourselves and our repositories in the most genuine and sincere way possible: self-knowledge and empathy" (144). It is with this affective approach that the archivist as an individual rather than a monolithic institution can make more meaningful connections with donors and users on a personalized level.

For example, the archive of a small theatre based in New York City only has one archivist on staff. Their user community includes academics and students, but also past and present artists performing at the theatre, visiting artists traveling from other states or countries, and the programming and marketing teams working for the theatre. It is a small community, many of the users know each other and know the institution well, and have very specific needs. What an artist might need from an archive, can be very different from the needs of the historian. The archivist in this situation may consider workshops designed to explore how artists can activate archives for use in their own work, rather than more traditional reference interviews, designed for academics. Or it may be more important for the archivist to prioritize access and discovery of their audiovisual collections if their user base is more interested in performance footage, rather than focusing on broad access to the whole collection. The archivist at the small institution can consider far more specific user needs, and decide what access looks like for their community, rather than generalizing for a wider population.

Additionally, archivists can gain knowledge of collections at a level not attainable in larger archives, where again, roles are often more specialized and work is divided amongst numerous staff members. Only one archivist working on a collection during accessioning, processing, digitization, and other processes means a greater level of consistency in arrangement and description, which can in turn create a more consistent access and discovery experience for patrons. Perhaps most importantly, it gives the archivist intensive time with each collection, allowing them to gain a deep understanding of the content and context of the materials, making the archivist more knowledgeable and effective when doing reference work and outreach utilizing those collections.

Leveraging their combined in-depth knowledge of collections and users' needs, archivists in small archives can work to create a system of discovery and delivery that sets realistic goals, better meets researchers where they are, and reflects the specific archive, as opposed to reaching for broader expectations of archival discovery and access systems. Smaller archives also can create an approachable environment where fewer barriers exist to accessing records and having memorable encounters with collections. Archives that center their community first may choose to "revamp arrangement and description both to represent their materials and to provide for findability, access, and (re)use on — and often literally in — [the community's] own terms" (Poole 2020, 663), rejecting larger institutional norms in favor of local practices that build trust and respond directly to community needs. Creating systems of access and discovery in dialogue with the archives' community allows for greater responsiveness to user and donor needs while also creating a sense of ownership for collection stakeholders.

In building relationships with their community advocates, smaller archival institutions may be better situated to communicate the procedures and policies involved in making collections accessible both physically and digitally. Additionally, they are well poised to seek input from donors and patrons when

determining and implementing strategies for increased access to collections because of their close community ties. This open line of communication can also provide valuable insights when determining what resources or systems are needed to increase access. Conversations between archivist and donor regarding rights management, user needs, policies, and procedures (amongst other subjects) are more manageable when a close relationship between archivist and donor exists.

It then follows that a smaller archive can create a more specific focus on what activities and collections the archive will prioritize when developing access systems, and can then create internal processes and procedures that fit best for the priorities they set. Policy and procedure development can be more streamlined - but this work must still take into account the needs of everyone who will be impacted by the work being done. While large repositories may need to reach for idealized “best practices,” smaller organizations with limited resources must assess how they will best be able to carry out archival work to their own internally-determined abilities. Having a smaller decision-making team for setting responsibilities can bolster the achievability of any given archival process because the work does not need to apply across many different employees or departments.

This flexibility of internal decision-making and stronger understanding of their collections and user community gives a small archive the ability to set focused priorities that will best serve those collections and users. Specific to access, small archives have advantages in setting procedures for processing collections, understanding preservation needs, determining access and handling guidelines, and clearly delineating all aspects of a collection management policy. While larger organizations may feel pulled in many different directions, smaller archives can create and implement more specific policies, which can better support the development of workflow priorities. As MacFarland describes, “Small archival programs are uniquely positioned to be innovators and leaders in this area [of focusing on archival users and staying positive under adverse conditions]. Because lone arrangers are providers of both public and technical services, we can focus on users by reallocating our time and attention” (147), or put another way, by being flexible and reprioritizing certain expectations.

Self-Assessment Discussion

There is often an assumption that if an archive is small and lacking in robust online access, that it must be undiscovered and therefore unused. But archival access exists on a wide spectrum, and should be treated as an active process rather than a set goal. It also takes many forms outside of the finding aid, online discovery systems, or the regular appearance in the footnotes of books and articles. In this sense, a small archive may not have a global reach, but that certainly does not mean it exists without a community of patrons and users. The house museum, the local historical society, the theatre, the community college – all of these institutions often have archives with communities both internal and external that make regular use of their collections. In the words of archivist Meredith Evans (2015), small archives should have room to “...[change] the definition of a successful archival program, special collections department, or comprehensive collection. Decide on realistic goals and consult with your administration or your advocates/allies who can help accomplish the goals to successfully develop, maintain, and ensure access to these collections” (11). In an effort to shift away from a culture of comparison, unrealistic expectations, and a scarcity mindset, archivists working at small institutions should feel encouraged to look first toward what they *do* have, rather than what they lack. The small scale of operation is the source of many of the challenges we’ve identified, as well as the source of many strengths and opportunities.

With this in mind, the following questions are provided for consideration. They have been developed with a focus on the lone arranger, but they are relevant to archivists working at institutions of all sizes. These questions are offered as a way to reframe thinking around what access and discovery can look like, how these two integral aspects of archival work can be improved, and what it means to be successful within the particular context of your archive.

- **Who is your user base/what communities are you serving?** Are you and your administrators on the same page about who this community is? How familiar are you with them, and with their archival needs? As discussed, having a clear understanding of the user community that engages with your archive is essential. Formalized in the DACS “Principles of Archival Description” (SAA TS-DACS 2020), the second Principle supports the importance of this work for each archive by stating, “Users are the fundamental reason for archival description... To make wise choices about descriptive practices [that impact how users find, identify, select, and use archival records], archivists must develop and maintain an awareness of user needs and behaviors” (xiv). Do users (or potential users) know how to navigate an archive, or your discovery process in particular? Consider building out documentation of any current discovery and delivery systems that you have in place. What projects, initiatives, or technology would improve community relationships? Do all current projects, initiatives, or technology directly serve the research needs of this community (and can you downscale what doesn’t)? Focusing on your user community will help support all other priorities.
- **How do you define access?** What actions are you taking daily that could be considered as contributing to improved access? Regular communication with your community of users, clear policy, public programming, and developing curriculum all contribute to improved access. How do you welcome people into your archive? Access is often discussed as the ability of a patron to use material physically or virtually, but improving access also involves empowering users to see the collections as within their reach. In other words, do users feel as though they have the right to use the materials? What kind of gatekeeping, unintentional or otherwise, might be preventing people from using your collections? What outreach have you done to reach both internal and external users?
- **What institutional knowledge do you have that isn’t documented?** How can you prepare others to continue the work you’ve put in place? What other policies and procedures are needed for supporting the sustained management of this archive? There is no such thing as “too much documentation” in the world of the archivist. The trick is keeping it all in one location! Policies should be written in order to meet specific internal needs, and documentation should be in plain language for non-archivists. While policy development should generally include a review and understanding of field standards, rather than borrowing complete policies from other institutions, it is essential to develop institutional policies with language specific to your archive. Policies require endorsement from the organizations’ governance, such as a Director or Board of Trustees, but that process can be a strategic opportunity to raise awareness of the functions and activities of the archive. Procedures, on the other hand, detail the “how to” aspect of accomplishing the work described in a policy, and they state the action steps that will be taken. Procedures and guidelines can be more informal, but should always be thoroughly documented.

- **What do you see as the expectations of your institution and your user base?** Are those expectations realistic and formed with a full understanding of the work you do? Rather than internalizing this pressure to deliver at a scale and speed outside of your archive's reach, communicate your priorities to administrators, and work together to set appropriate expectations. Make these expectations transparent and accessible through written policies and procedures so external users can see the labor involved in their requests and know what to expect.
- **Who within your institution do you need to connect with in order to ensure the archive is included in institution-wide strategic planning?** How can you ensure that the needs and goals of the archive are included in the conversation? Bring invested parties into the planning and development stages of your projects to ensure continuous communication between you and your administrators. Find additional ways to demonstrate the value of the archive to unengaged administrators. Do you have an elevator pitch? What partnerships can you build to introduce the archive to your internal community? For example, could you partner with HR to promote the archive at new employee orientation?
- **What is your archive's mission and what are the parameters/boundaries of your position?** Are these goals and expectations clear to your administrators? Ensure that the archive's mission is in line with (and included in) the institutional mission. Document the requirements and expectations of your position through a clearly written job description that is reviewed regularly. Having these statements in writing will provide a clear framework to return to when needed to support project, staffing, and funding requests by demonstrating how these needs are integral to the mission and work of the archive.
- **When planning new projects, consider what is the true cost of new technology?** Because an application is inexpensive, or even open-source, it doesn't mean that budget considerations should stop there. Server and application maintenance can be a significant time and monetary investment. Local IT may or may not have the bandwidth to maintain this for your archive, and vendors, while expensive, may bring peace of mind (and potentially represent cost-savings to administrators who also budget for IT). Allow room for the possibility that the challenges you face won't be solved by finding the perfect piece of software, and that a less-is-more approach may be more sustainable.
- **Are you looking to other institutions to serve as a model? Do those institutions serve as an appropriate model?** It's easy to point to well-resourced institutions with robust technology landscapes, but the cost and manpower required to maintain them may not feel achievable to a small archive. Best practices as defined by the wider profession may not always fit the resources of a small institution. Rather than treating those best practices as the bar you must meet in order to be considered successful, instead treat them as a framework to be customized to suit the specific needs and realities of your archive. Consider looking to peer organizations and community-led archives for more approachable models of archival practices.
- **Are inter-institutional or inter-departmental collaborations available and advantageous?** These types of collaborations can be a great way to adopt established technologies and workflows, while promoting the visibility of your collections by linking them to other departments or institutions. However, if these technologies or workflows are not well-suited to

your archive's needs and available resources, it may ultimately be more effort than is worthwhile. A collaborative project should be one that is truly mutually beneficial, leaving your archive with a higher degree of visibility *and* providing long-term sustainability for the project.

- **Is this project too big?** Are the outcomes of this project sustainable in the long-term? This perhaps speaks for itself, but is a valuable question to ask at the beginning of any grant application or proposal. It is okay to operate within the scale of your institution, and in fact will make for a richer, more successful project. This can occasionally be a hard line to hold if you are facing pressure from leadership or administrators to take on a larger, attention-grabbing project, but make your case in terms of what is sustainable after the initial grant period is over.
- **How are you measuring success?** Do you need to set more scalable goals? Remember, what success looks like at a large, well-resourced institution will not be what success looks like for a small archive. Set goals that will help you reach measurable improvements and outcomes for your individual archive. Setting those goals in collaboration with your administration and leadership will ensure you have buy-in and support when it comes time to make important decisions. Archival work is inherently iterative and long-term in nature, so use this to your advantage in advocacy and planning work.

This set of questions is not meant to be comprehensive or complete, but rather to serve as a jumping-off point into the more specific strengths of your archive. While small archives tend to be perpetually underfunded and understaffed, it is often impossible for the solo archivist to solve these challenges entirely on their own. Instead, use these questions to begin assessing where these issues currently stand in your archive. Use the answers to help begin to plan small, achievable steps forward. Include internal stakeholders as well as your community and patrons where it feels relevant and helpful. Having administrative buy-in and community support can go a long way in helping to achieve those next goals. Rather than feeling immobilized by what is out of reach and allowing “perfect to be the enemy of good,” focus instead on what you can accomplish and on achievable goals for your unique archive.

Acknowledgements

The authors would like to acknowledge Linda Hocking, Archivist for the Litchfield Historical Society, and the facilitator for our Lighting the Way working group. Her insight was invaluable and greatly appreciated. We thank her for all of the time she spent working with us throughout the project, from our first meeting through the final draft of this paper.

References

- Buchanan, Sarah, Jane Gruning, Ayse Gursoy, and Lecia Barker. 2017. "Surveying Archivists and Their Work toward Advocacy and Management, or 'Enterprise Archiving.'" *The American Archivist* 80, no. 2 (Fall/Winter): 268–95. <https://doi.org/10.17723/0360-9081-80.2.268>.
- Evans, Meredith R. 2015. "Modern Special Collections: Embracing the Future While Taking Care of the Past." *New Review of Academic Librarianship* 21 (2): 116–28. <https://doi.org/10.1080/13614533.2015.1040926>.
- McFarland, Colleen. 2007. "Rethinking the Business of Small Archives." *Archival Issues* 31 (2): 137–49. <http://digital.library.wisc.edu/1793/72281>.
- Matienzo, Mark A., Dinah Handel, Josh Schneider, and Camille Villa. 2020. *Lighting the Way: A Preliminary Report on the National Forum on Archival Discovery and Delivery*. Stanford University Libraries. <https://purl.stanford.edu/rt078dm2196>.
- Poole, Alex H. 2020. "The Information Work of Community Archives: A Systematic Literature Review." *Journal of Documentation* 76 (3): 657–87. <https://doi.org/10.1108/jd-07-2019-0140>.
- Society of American Archivists Technical Subcommittee on Describing Archives: A Content Standard (SAA TS-DACS). 2020. *Describing Archives: A Content Standard (DACS)*. Version 2019.0.3. Chicago: Society of American Archivists. https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf.

Additional Readings

- Casari, William. 2015. "Community Archives and Autonomy: Meeting the Challenge to Preserve Vital Records of the College and the Community." *Journal of the Society of North Carolina Archivists* 12: 19–40. http://www.ncarchivists.org/wp-content/uploads/2017/08/JSNCA_2015_Vol12_Casari.pdf
- Erway, Ricky. 2012. "Increasing Access to Special Collections." *LIBER Quarterly* 21, no. 2 (January): 294–307. <https://doi.org/10.18352/lq.8026>.
- Gasero, Russell L., Chana Kotzin, Lisa M Sjoberg, and Alison Stankrauff. 2011. "Which Hat Are You Wearing: 'You Need What? When?'" (Session 208). *The American Archivist* 74, Supplement 1: 1–32. <https://doi.org/10.17723/aarc.74.suppl-1.l40119167t266160>.
- Haberstock, Lauren. 2019. "Participatory Description: Decolonizing Descriptive Methodologies in Archives." *Archival Science* 20 (2): 125–38. <https://doi.org/10.1007/s10502-019-09328-6>.
- Heslin, Kayla. 2019. "Exploring Rockingham County's Past: Recapturing Local History and Promoting Accessibility." *Journal of Contemporary Archival Studies* 6, Article 29. <https://elischolar.library.yale.edu/jcas/vol6/iss1/29>.

- McCarthy, Jack, Celia Caust-Ellenbogen, and Sarah Leu. 2016. "Building an Inclusive Community of Archival Practice: The Historical Society of Pennsylvania's Hidden Collections Initiative for Pennsylvania Small Archival Repositories." *Pennsylvania History: A Journal of Mid-Atlantic Studies* 83 (1): 97–102. <https://doi.org/10.5325/pennhistory.83.1.0097>.
- Trammell, Anna. 2020. "Taking the Plunge: When Improving Access and Discoverability Means Walking Away from Your Existing Systems and Starting Over." *Lighting the Way: A National Forum on Archival Discovery and Delivery*. Presentation, February 10. <https://purl.stanford.edu/pr103tf2767>.
- Wickham, M. Sarah. 2015. "Heritage Quay: What Will You Discover? Transforming the Archives of the University of Huddersfield, Yorkshire, UK." *New Review of Academic Librarianship* 21: 195–205. <https://doi.org/10.1080/13614533.2015.1042116>.
- Zamon, Christina. 2013. *The Lone Arranger: Succeeding in a Small Repository*. Chicago: American Library Association.

Section 3: Emergent Opportunities

The Power of Parallel Description: Wikidata and Archival Discovery

Kelli Babcock, Regine Heberlein, Anna Björnsson McCormick,
Elizabeth Russey Roke, Greta Kuriger Suiter, and Ruth Kitchin Tillman

Abstract: This paper proposes the use of Wikidata in archival description workflows to create linked data that can improve discovery of our collections and expand contextual information around the entities they represent. Wikidata offers an opportunity for opening up archival description to a public forum. We suggest that, rather than approach it as a means to drive traffic to our individual finding aids, archivists use it as a platform to transform information about entities already represented in our archival description into open data. In turn, other communities can engage with these contributions to enhance, redescribe, or replace oppressive or harmful language—with all of the potential for collaboration, as well as conflict, that this can entail. Community description is a fundamental departure from traditional descriptive practice. It may significantly shift descriptive work for archivists, and it requires engaging with the Wikidata community to understand the methods and purpose of the infrastructure. With so much to consider when assessing whether an archives should use Wikidata, the paper is supplemented by an actionable list of 50 items that archivists can consider to get started.

Setting Context / Wikidata's Relevance to Archival Description

Wikipedia went live on January 15, 2001, and quickly became one of the most popular sites on the internet. Wikimedia projects rapidly expanded to include new sites for the sharing of public domain materials. These include Wikisource, a collection of public domain and freely-licensed texts, and Wikimedia Commons, a repository for all of the images used in Wikipedia. In October of 2012, Wikidata joined the Wikimedia projects as an open, editable knowledge base consisting of structured and unstructured information. Initially, its data was derived from other Wikimedia sites, but the project invited anyone to expand its coverage through adding or enhancing records. Today, there are over 93 million items in Wikidata.

Recent collaborative efforts in understanding and utilizing Wikidata in the GLAM community (Galleries, Libraries, Archives, Museums) include the LD4 Wikidata Affinity Group and the Program for Cooperative Cataloging (PCC) Wikidata Pilot. Both groups include members from across the GLAM community, but most are from higher education institutions. These projects bring together librarians, catalogers, archivists, metadata experts, and IT professionals among others. Even though these individuals often work in the same institutions, they usually approach metadata and description projects from different points of view and, often, different standards. Some work at the item level (books, articles, images), others work at the aggregate level (collections of items), and some are used to manipulating data using

APIs and scripts. The technical systems that GLAM employees use to keep track of collections vary between institutions and according to budget and staff resources.

This paper considers the issues surrounding archival description and Wikidata. Contributing to Wikidata and helping to shape Wikidata is potentially a powerful opportunity for GLAM institutions that will enable us to expand the reach of archival description, attract new users, and contribute knowledge to Wikidata. This effort, however, should be approached carefully and thoughtfully. Each community that interacts with the Wikidata infrastructure will bring its own set of goals, biases, standards, and ethics. Understanding and supporting the Wikidata ecosystem is just as important as achieving the desired outcomes and contributions of one's own community. We must participate as members of the Wikidata community, not colonize and exploit it to our own ends.

The Wikimedia Foundation strategic direction states: "By 2030, Wikimedia will become the essential infrastructure of the ecosystem of free knowledge, and anyone who shares our vision will be able to join us." For GLAM purposes, "join" is the operative word in this statement. Those who use, edit, contribute to policy, or otherwise interact with the Wikimedia community have an opportunity to engage, and become a part of, a dedicated group of volunteers who care passionately about organizing and presenting information to a global audience for free, available both online and off. There is no other organization in the world that brings together as many people around a shared vision that does as much public good.

When it comes to shared values, the PCC, Wikimedia Foundation, and the Society of American Archivists have much in common. All three organizations firmly state that their values include: innovation, collaboration, diversity and inclusion, and sharing expertise (Program for Cooperative Cataloging, n.d.; Wikimedia Foundation, n.d.; Society of American Archivists, 2020). Wikimedia projects excel at creating platforms for all to edit. Libraries and archives can greatly benefit from Wikimedia's infrastructure, but those working in GLAM institutions need to also bring awareness and respect to how they engage with Wikimedia projects. Some questions to consider before editing Wikimedia projects include: What rules guide conduct in this space? How might our involvement in Wikidata benefit both the Wikidata and the archival community? How can GLAM organizations join and utilize this infrastructure? Where does Wikidata fit into the universe of archival discovery and description? Where does archival discovery and description fit into the universe of Wikidata?

The recommendations that follow are an attempt to outline the major benefits and roadblocks for archives looking to engage in the Wikidata community.

Wikidata for Linked Data

Implementing linked data sustainably has proven elusive for the GLAM community. The Technical Subcommittee on Encoded Archival Standards (EAS) has gestured towards the efficiencies of linked data by breaking out select entities, most notably corporate bodies-persons-families (EAC-CPF) and functions (the proposed but inactive EAC-F), into their own encoding standards separate from the description of records (EAD) and providing for linking between them. Social Networks and Archival Context (SNAC), a cooperative project that aggregates information about "persons, families, and organizations that created or are documented in historical resources (primary source documents) and

their connections to one another" uses EAC-CPF files to build its database (SNAC Cooperative, n.d.). However, the EAC-* standards remain far from a true linked data model.

FRBROO, the first object-oriented conceptual model attempting to bridge the library and museum domains, was published in 2006 and has been updated several times since. However, lacking infrastructure and hampered by its considerable complexity, it has been implemented by only a handful of projects (e.g., the DOREMUS project) (Doremus project, n.d.). Records in Contexts (RiC), a conceptual model and ontology developed specifically for archival description that has been under development since 2012, only released a stable version of its ontology earlier in 2021; the ontology remains in draft as of this writing. Next year, 2022, marks 10 years since BIBFRAME, one of the first linked data ontologies built specifically for libraries, was first released, but its use in production is limited to a few institutions.

For the vast majority of cultural heritage institutions, their only experience with linked data is experimental, if they engage with it at all. Building new linked data infrastructure and systems requires dedicating resources not just to implementation but also to their long-term maintenance. The lack of an out-of-the-box solution makes linked data implementation a heavy lift for all but the most well-resourced institutions that have technical expertise, time, and resources to build new platforms, systems, vocabularies, and ontologies, as well as staff time for skill-building and experimentation. At a time when budgets in many cultural heritage institutions continue to decrease and we are asked to do more with less, building a new metadata infrastructure for linked data may not be a sustainable option nor a priority for many institutions.

Wikidata removes some of these barriers to implementation. As demonstrated in OCLC's Project Passage (Godby et al., 2019), the PCC Wikidata Pilot, and further explored by the OCLC Research Archives and Special Collections Linked Data Review Group (2020), Wikidata (and the Wikibase software underlying it) offers viable technological solutions for institutions looking for an accessible entry point into linked data that is relatively low-cost, doesn't require building new technical platforms from scratch, and isn't dependent upon yet-to-be completed domain-specific ontologies. While this does not mitigate the need for skill-building among staff and the development of new descriptive workflows, Wikidata has the potential to dramatically lower the bar for linked data, providing an intuitive and easy-to-use editing interface, automated identifier creation, data publication, and out-of-the-box querying. Additionally, Wikidata can serve as a linking hub for matching identifiers from over 1300 different systems together in one place (van Veen, 2019). With Wikidata anyone can create linked data, requiring only internet access and a free user account. Wikidata also has tools for both individual item creation and batch loading. As a technological solution, Wikidata (and the open-source Wikibase platform that supports Wikidata) show great potential for the archival descriptive community. It is a scalable, plug-'n-play solution that can meet many of the linked data needs of archival repositories, from large institutions to lone arrangers.

Although these platforms remove some technological roadblocks for the archival community, it is important to recognize that it is not a magical solution and there are limitations to its use. While Wikidata does have some batch loading tools, they can be cumbersome to use. Some archival entities may not lend themselves to batch loads, requiring inefficient and potentially unsustainable hand-editing of entities. Other concerns are external. A primary concern is the funding model that underlies Wikidata. Wikidata is supported by the non-profit Wikimedia Foundation, dependent upon fundraising and grants for its existence. Although Wikimedia just celebrated its 20th anniversary, there is no guarantee that Wikidata and, most importantly, the data it contains, will exist in perpetuity.

As no information in Wikidata is institutionally owned or tracked like in traditional library consortial systems, it is unclear how or if it would be possible to remove data efficiently. Wikidata is also community-based, open to anyone who wishes to create an account. While this can be a potential benefit of the tool for archives looking to expand community involvement in their collections, constantly changing data, inaccurate data, or worse, data that has been maliciously altered are all very real possibilities. As participants in the PCC Wikidata project have noticed, bots created to add identifiers to entities, for instance, sometimes add the wrong identifier to entities creating headaches for those using Wikidata for matching procedures.

Wikidata Opens Description to Multiple Communities

Moving Towards Community Description

Wikidata positions its purpose and mission outside the traditionally exclusive realm of professionalized description as “a free, collaborative, multilingual, secondary database, collecting structured data to provide support for Wikipedia, Wikimedia Commons, the other wikis of the Wikimedia movement, and to anyone in the world.” (Wikidata:Introduction 2021) Many of those statements stand in stark contrast with current practice in archival—and, more broadly, GLAM—resource description. Whereas the creation of descriptive metadata is a professionalized activity in GLAM circles, with practitioners undergoing years of training in the application of domain-specific standards, Wikidata describes itself as both *collaborative* and *open to anyone in the world*: “data is entered and maintained by Wikidata editors, who decide on the rules of content creation and management.... Anyone can use Wikidata for any number of different ways” (Wikidata:Introduction 2021).

Adopting Wikidata for GLAM resource description, therefore, represents a significant departure from previous descriptive practice. For a gated community operating in a space of professional standards and professional “codes” of implicit and explicit expectations, what are the implications of using radically open description—description open to “anyone in the world”; open to non-professionals and professionals alike; open to contributors from different, even non-compatible domains; open to the human spectrum of expertise, intention, and ideology; open to contributors even at cross-purposes?

Records in Contexts (RiC) takes a clear position on that question. In their preamble entitled “Why RiC-O?”, the authors answer unequivocally that “relying on existing ontologies outside of the control of the archival community presents risks, as any one of these ontologies may change the meaning or scope of any given concept at any time, rendering its use in archival description inaccurate.” (RiC-O, n.d.) In a context of professionalized resource description, the assessment that giving up control over description incurs the risk of destabilizing meaning is as insightful as it is consequential, and its underpinnings may bear some closer examination. What would happen if archivists ceded control?

Archivists and other GLAM professionals operate within a code of standards that tends to be tacitly exclusive and monolithic. It is not uncommon for materials to be dispersed among museums, libraries, and archives, respectively, according to format or, simply, the mechanics of acquisition. The documents from a university’s archaeological expedition might be sent to the archives, the pottery to the museum, the coins to the Numismatics collection. A provenance-based collection might be parceled out to multiple repositories by the donor or by the way the material is bundled into lots for sale. Subject-based

collections are rarely in communication with one another across institutional boundaries other than through cross-references. These realities, compounded by the discursive legacy of the finding aid (Wiedeman, 2019), are apt to create encapsulated collections, represented within the confines of narrative finding aids, that are hard for researchers to situate in cross-institutional contexts. In this way, GLAM description in its current state frequently privileges one single context and narrative over others.

Similarly, the descriptive standards used by the community, while meant to facilitate discovery and access by ensuring uniformity, can do the opposite by imposing one voice and one judgement over others under the mantle of professional authority. Eliminating bias in libraries means re-examining authority records, subject headings, Library of Congress call numbers, and description in finding aids as well as the descriptive practices that created them. Gender, race, immigration status, colonialism, or religion in authority records and descriptions are among the examples the GLAM community has already begun to discuss that explore the potential of linked data to create more equitable solutions (Allison-Cassin and Scott 2019; Douglas et al. 2018).

Wikidata presents an opportunity for GLAM institutions to challenge these practices. By providing archivists as well as their stakeholders with an open linked data platform, it allows archivists to cede control over description and invite community participation. It provides archivists an opportunity to address harmful labels or subject terms, assigned without consent by the communities they are intended to describe. At the same time, community description can open the door to errors, opposing viewpoints, even malicious falsehoods. One contributor's truth may be another's deceit. Properties such as gender in Wikidata highlight this tension (Wikidata:WikiProject LGBT/gender 2021). Yet with equitable description at stake, this is perhaps a risk worth taking. One might say the risk of destabilizing received knowledge is the point of using Wikidata for GLAM description. No single statement about a Wikidata Item stands alone or uncontested; it can be improved upon, augmented, edited, corrected, or even contradicted by anyone. The resulting metadata is not resource description as we know it. It no longer speaks with one voice. It no longer speaks with unfettered authority. It is apt to create new and polyvalent meanings around records—constantly, insistently, loudly. It is messy, inconsistent, and ever-evolving as humans themselves.

In this, Wikidata can be used to create anti-oppressive, reparative, decolonized, inclusive archival descriptions and discovery. It can counter bias, misinformation, or error by providing supplemental metadata created by anyone. The case study "Using the Web of Data to Study Gender Differences in Online Knowledge Sources: The Case of the European Parliament" (Hollink et al., 2018) found less gender bias in Wikidata items representing members of the European Parliament than other sources. The study found little difference between male and female Wikidata entries for EU parliament members and states, "...it is possible that a subtle bias is noticeable in natural language text of Wikipedia articles but not in the structured data that we analyzed." Since GLAM institutions mostly describe in narrative ways, moving to community description as facilitated by a linked data approach and mindset could lessen inherent bias in descriptions of people.

In such an ecosystem of linked open data, archivists may be called upon to consider fundamental questions about their stakeholders and the focus of archival work, with the inevitable long tail of questions of how archival work is taught in academic programs and how it is evaluated and rewarded within institutional promotional structures. More likely than not, committing to linked open data will bring about a shift in priorities away from physical processing tasks towards more abstract activities

such as conceptual modeling, ontology work, and ethical and legal questions. In the course of this shift, some fundamental questions the community may need to reflect on include:

- What are the values we implement when creating data in an open data environment where our practices may not always dovetail comfortably with others’?
- To what level of generality should description adhere to serve an open community; and to what level of specificity—including professional domain-specificity—should description go to serve special-interest communities? Where do archivists localize themselves and the communities they serve on that spectrum?
- Can we get comfortable with creating description and discovery in a system that allows for multiple ontologies that may not be rooted in professional standards?
- What criteria will we apply when interacting with the contributions of others, e.g. in the form of harvesting metadata into our local environments or, conversely, overlaying open data with edits of our own?

Practicalities and Parallel Descriptions

To best understand how Wikidata can be used to enhance cultural heritage description and discovery, it is important to understand the current descriptive practices of librarians, archivists, and other GLAM workers and explore how Wikidata can be integrated with the work we are already doing. Archivists and librarians have historically siloed our descriptive data. Intellectually, we create metadata according to our own content standards and use our own judgment to determine whether a piece of information about a collection or entity is worth including. Technologically, we have developed a few sites of shared descriptive records but still keep much of it in our own systems or on our websites.

Almost all repositories use some kind of software as a system of record for their repository’s holdings as well as to manage their collections. This data could be managed using a spreadsheet or simple database, but many institutions use information management systems built for archives such as ArchivesSpace and Access to Memory (AtoM). Archivists record data in a structured way by following descriptive standards such as Describing Archives: A Content (DACS), General International Standard Archival Description (ISAD(G)), or the Rules for Archival Description (RAD), and publish this data as finding aids. While archival description is focused on describing entities such as records and their various aggregates, in the course of writing description archivists also describe the entities that created, donated, or are otherwise represented in the records. Archivists describe entities with narrative description; in some cases (e.g. people, organizations, or places) authority records such as the Library of Congress Name Authority Files are referenced.

In short, descriptions are the labor of archivists that enable collections to be found by researchers, but they are structured in ways that require description creators to have specialized education and knowledge of bibliographic and archival systems and standards. Potential researchers must also know where and how to look for the information they are seeking. The end result is that the description we create for our collections is often siloed and can be difficult to locate outside of academic search tools.

By providing a plethora of descriptive properties which can be used in any (technically valid) combination and repeated, Wikidata blows these descriptive practices wide open, resulting in sets of *parallel descriptions* where statements associated with multiple domains, standards, or data models are allowed to be attached to the same entity. In contrast to the highly domain-specific data models currently employed by the GLAM community, this feature facilitates the repurposing of existing Wikidata items across domains regardless of their author or their intended domain context at the time of creation.

In allowing for the mapping of domain-specific concepts to Wikidata as Wikidata Items—and for the subsequent linking of other Wikidata Items to them using the “instance of” property—Wikidata also offers the option of a meta-layer of sorts of parallel description: a layer of properties that make it possible to explicitly retain, where desired, a domain-specific conceptual model. For example, a given archival resource may be described in parallel as an “instance of” a “fonds” (Q3052382: “aggregation of works that originate from the same source”) as well as an “instance of” a “record group” (Q59294700: “in archival science, the largest level of description; the whole of a collection established according to provenance, such as the records related to an entire organization”); the two concepts may not mesh perfectly nor serve every user, but they can support domain-specific querying when applied.

The consequences from allowing parallel descriptions to co-exist, even contradict one another are as simple as they are revolutionary. Wikidata Items can be used as general-purpose entities and linked to one another to form statements. This means Items created by others on Wikidata can be re-purposed by anyone for any statement, releasing data creators from the need to re-create an Item in domain-specific ways (and then clumsily chaining them together with properties indicating sameness). Items so repurposed can be used to link across the boundaries of institutions or national borders, uniting, e.g., collection objects that share a context but are physically split between museums and archives, or dispersed among repositories worldwide. Being linked on one platform across the boundaries of format, custody, ownership, or professional domain, items can form part, at one time, of multiple contexts and multiple conversations. For example, current work underway at the Jewish Digital Cultural Recovery Project (JDCRP) uses an (as yet private) Wikibase implementation to trace the ownership of lost and stolen artworks over time, representing provenance evidence in one descriptive network (Jewish Digital Cultural Recovery Project, n.d.).

There is no question that this model adds significant complexity to current descriptive practice and destabilizes long-held foundational notions about “records” and “collections”. In Wikidata, an Item or entity page is not a “metadata record” in the traditional library and archival sense. Rather, it is a set of statements that can be combined to create a general understanding of the entity as described by those who contributed. Likewise, the boundaries of a “collection” on Wikidata expand with each intersecting contextualization. Because there is no standard for completeness in Wikidata, each contributor works from their own knowledge, resources, and priorities, or those of their community. Yet despite a lack of consistency across entities, we can benefit significantly from this prior work. The relationships encoded in these statements create a network of links which can be followed to construct contexts. For example, the query which could be expressed as

SELECT [person] and [archival repository] WHERE the person was educated at an institution, the institution is an HBCU, and their archives are at the archival repository

yields 180 results at time of writing (see <https://w.wiki/3V3F> to run this SPARQL query). These results are dependent on someone having added the following things (and more) to Wikidata:

- an entity representing the person,
- a statement that the person was educated at an institution or institutions,
- a statement "is instance of HBCU" to the educational institution's entity, an entity representing the archival repository,
- a statement that HBCUs are a subclass of institutions of higher education,
- and a statement that the person's archives are at the repository.

The more our own data is represented within Wikidata, the more we can pursue such cross-institutional inquiry. Each of these entities and statements could have been made by one person but were more likely made by five people, only one of whom was in the archival field. And because records are never complete, there are likely other less famous graduates of HBCUs who can be found in our repositories but who wouldn't appear in this query because they're missing the "educated at" or "archives at" statements on their records.

For those accustomed to control of our data, using incomplete data which could have been created by anyone may be a challenge. Whenever we engage in open platforms, we worry about control, quality of data, and validity (Seeman and Dean 2019). Wikidata properties describing an entity may repeat and even conflict with each other. While troubling to those of us who seek a single "right" answer, the same errors, disagreements, biases, and priorities are present in the primary and secondary sources we use for our own description, in the records we create, and even in the systems we use. An example of this is incorporating undated encyclopedia entries alongside search results in discovery systems like Summon (Tillman 2021).

Where, then, do our standards play a role? Content standards, in particular, provide a starting point for us to ask "what can we contribute" and "what will be of use." Three members of our group had been part of an independent Archives and Linked Data Interest Group which created recommendations by mapping archival content standards to Wikidata properties. (Wikidata:WikiProject Archives 2021) When developing the recommendations, they considered the function of each content element and its appropriateness to Wikidata. A biographical and historical note might be DACS "Optimum", but Wikidata's overall community does not benefit from paragraphs of text, nor does it support functional querying as above. Instead, one could extract specific relational statements from such notes and encode those instead, using appropriate properties. While extracting these relationships from existing textual notes is an onerous proposition (see OCLC's archives and special collections linked data research on navigating between notes and nodes), a worker creating new description must already identify these relationships and could add them to Wikidata (or discover them in Wikidata) in the process of writing such notes (OCLC 2020).

Editing Wikidata also allows us to serve as a way to direct people to our more conventional description and puts our collections in context with other resources outside of our own repositories. Items owned by different collections can be linked together across institutions or countries as well, freeing our

description from the silos to which it is often relegated. The “archives at” property can be added as a dynamic infobox to the entity’s Wikipedia page in addition to an External link. This can drive referral traffic to our institutional discovery platforms while simultaneously contributing to open knowledge on the web.

While our discovery platforms may be optimized for search engine indexing with sitemaps and other SEO features, creating parallel descriptions in an open tool such as Wikidata allows description data to be a part of the web. The data is already structured and available to be pushed from our siloed databases into Wikidata. Wikidata also allows for each institution to create linked open data within their own capacity constraints. One institution may choose to create full parallel descriptions, while another may simply add an “archives at” property link to their internal finding aid description. Both approaches benefit the user by locating an entity’s archival records, and also potentially added new contexts and relationships around that entity.

Wikidata is an alternative option for providing access to our archival records, and comparatively fluid when considering the rigid standards of traditional archival description practices. While standards serve a purpose, creating parallel descriptions in Wikidata allows us to operate in a much more flexible and collaborative information environment. Wikidata eliminates the gatekeeping associated with traditional modes of description and discovery: it is free, open to everyone, and there is the ability to add as much or as little information as you want. A Wikidata item is never complete or incomplete, it just is. As such, it allows us to dip our toes into the Wiki* universe without feeling pressure to do everything at once. We can begin by doing what we are able to do and iterating on our efforts as time goes on.

How Do We Move Things Forward?

In each author’s institution, Wikidata was selected for exploration because it was accessible, free, and quick to adopt. Exploratory Wikidata projects were initiated simply by creating a Wikidata account and reviewing Wikidata documentation from other GLAM institutions. Wikidata as a tool also provided an intuitive web form interface to demonstrate linked data principles and how they could apply to the entities already existent in our descriptions. As our institutions began to use Wikidata, we witnessed how its function as an open knowledge base could append domain knowledge from contributors and users outside of our own institution to the items and properties we were building in Wikidata. Wikidata’s extensive list of external identifier standards, such as VIAF and SNAC, also offered the opportunity for our institutions to complete identifier reconciliation on internal authority records. We also noticed other institutions creating additional “archives at” properties to surface related records located at institutions that we had previously been unaware of. What other archival description platform can provide these same opportunities? Even if one could be funded and sustained, it would not come with the knowledge funneled into Wikidata by other domain experts working on non-GLAM projects in the platform. These features of Wikidata are not easily replicable in purpose-built archival discovery systems.

It should also be acknowledged that Wikidata does not eliminate all barriers. It preferences those with the time and money to contribute. It requires internet access as well as the ability to master a new and potentially complex user interface that is predominantly in English. These concerns remain unresolved and are an ongoing discussion in the Wikidata community and the library technology community

(Wikidata 2019; Matienzo 2015). We hope this paper is a small step forward in making Wikidata a more accessible option for archivists and GLAMs of all sizes as a tool for standalone description or supplemental discovery. To aid archivists considering Wikidata, we were inspired by the Collections As Data “50 Things” list (Padilla et al. 2018) and created a sister list to specifically identify [50 actions archivists can take to begin working with Wikidata](#).¹

Even as we offer Wikidata as a solution for building linked data parallel descriptions, we still have questions around this work. User experience research is needed to consider how Wikidata discovery tools can improve access and search for users. Assessment is needed to understand if, and how, external communities engage with the entities archives create. We must include our IT staff in explorations to engage with Wikidata, learn from it, and consider how it may be used to improve local discovery systems. We will continue to consider data stability in the context of a community-based system. While each author shared impressions of flexibility and efficiency when sharing their Wikidata experience during the Lighting the Way Working Meeting, all agreed data is required to demonstrate what a typical Wikidata “project” may include for an institution. Site analytics can be mined for referrals, and Wikipedia infobox work can be undertaken to bridge finding aid links to the creators who made the records. Archival description in Wikidata represents a fundamental shift in practice and an opportunity to make the changes we want to see in our day-to-day practice. As archival practice moves towards post-custodial and inclusive practice, it is up to us to re-examine old assumptions and look for solutions beyond our own walls.

Acknowledgements

The authors of this paper would like to thank Juliet Hardesty for her expert facilitation in conversations around this topic. The conclusions in this paper benefited greatly from her probing questions, constructive criticism, and patient guidance.

¹ A reference copy has been included as an appendix to this chapter.

References

- Allison-Cassin, Stacy, and Dan Scott. 2018. "Wikidata: A Platform for Your Library's Linked Open Data." *The Code4Lib Journal*, no. 40 (May 4). <https://journal.code4lib.org/articles/13424>.
- Doremus Project. n.d. "Doremus Ontology." Accessed July 1, 2021. <http://data.doremus.org/ontology/>.
- Douglas, Jennifer, Greg Bak, Evelyn McLellan, Seth van Hooland, and Raymond Frogner. 2018. "Decolonizing archival description: Can linked data help?" *Proceedings of the Association for Information Science and Technology* 55 (1), 669–672. <https://doi.org/10.1002/pra2.2018.14505501077>.
- Godby, Jean, Karen Smith-Yoshimura, Bruce Washburn, Kalan Knudson Davis, Karen Detling, Christine Fernsebner Eslao, Steven Folsom, et al. 2021. "Creating Library Linked Data with Wikibase: Lessons Learned from Project Passage." Dublin, OH: OCLC Research. <https://doi.org/10.25333/faq3-ax08>.
- Hollink, L., A. V. Aggelen, and J. V. Ossenbruggen. 2018. "Using the Web of Data to Study Gender Differences in Online Knowledge Sources: The Case of the European Parliament." *WebSci*. <https://doi.org/10.1145/3201064.3201108>.
- Jewish Digital Cultural Recovery Project. n.d. "About Us." Accessed July 1, 2021. <https://jdcrp.org/>.
- Matienzo, Mark A. 2015. "To Hell With Good Intentions: Linked Data, Community and the Power to Name." Presentation at the Library and Information Technology Association Forum, Minneapolis, MN. Accessed July 1, 2021. <http://matienzo.org/2016/to-hell-with-good-intentions/>.
- OCLC Research Archives and Special Collections Linked Data Review Group. 2020. "Archives and Special Collections Linked Data: Navigating between Notes and Nodes." Dublin, OH: OCLC Research. Accessed July 1, 2021. <https://doi.org/10.25333/4gtz-zd88>.
- Padilla, Thomas, Laurie Allen, Hannah Frost, Sarah Potvin, Elizabeth Russey Roke, and Stewart Varner. 2018. "Collections as Data 50 Things You Can Do." Accessed July 1, 2021. <https://doi.org/10.5281/zenodo.3066237>.
- Program for Cooperative Cataloging. n.d. "About the PCC." Accessed July 1, 2021. <https://www.loc.gov/aba/pcc/about/>.
- RiC-O. n.d. "Why use RiC-O?" Accessed July 1, 2021. <https://ica-egad.github.io/RiC-O/why-use-RiC-O.html>.
- Seeman, Dean, and Heather Dean. 2019. "Open Social Knowledge Creation and Library and Archival Metadata." *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 3 (February 27). <https://doi.org/10.5334/kula.51>.
- Society of American Archivists. 2020. "SAA Core Values Statement and Code of Ethics." Accessed July 1, 2021. <https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics>.

Society of American Archivists Technical Subcommittee on Describing Archives: A Content Standard (SAA TS-DACS). 2021. "Statement of Principles," *Describing Archives: A Content Standard*. Version 2021.0.0.2. https://saa-ts-dacs.github.io/dacs/04_statement_of_principles.html.

SNAC Cooperative. n.d. "About SNAC." Accessed July 1, 2021.
<https://portal.snaccooperative.org/about>.

Tillman, Ruth. 2021. "Time is a Property of Data." Accessed July 1, 2021.
<https://ruthillman.com/post/time-property-data/>.

van Veen, Theo. 2019. "Wikidata: From 'an' Identifier to 'the' Identifier." *Information Technology and Libraries* 38 (2): 72–81. <https://doi.org/10.6017/ital.v38i2.10886>.

Wiedeman, Gregory. 2019. "The Historical Hazards of Finding Aids." *The American Archivist* 82, no. 2 (Fall/Winter): 1-40. <https://doi.org/10.17723/aarc-82-02-20>.

Wikidata. 2019. "Wikidata:WikidataCon 2019/Program/Sessions/Keynote: Questioning Wikidata." Accessed July 1, 2021.
https://www.wikidata.org/wiki/Wikidata:WikidataCon_2019/Program/Sessions/Keynote:_Questioning_Wikidata.

Wikidata. 2021. "Wikidata:Introduction." Accessed July 1, 2021.
<https://www.wikidata.org/wiki/Wikidata:Introduction>.

Wikidata. 2021. "Wikidata:WikiProject Archives Linked Data Interest Group/Description Recommendations." Accessed July 1, 2021.
https://www.wikidata.org/wiki/Wikidata:WikiProject_Archives_Linked_Data_Interest_Group/Description_Recommendations.

Wikidata. 2021. "Wikidata:WikiProject LGBT/gender." Accessed July 1, 2021.
https://www.wikidata.org/wiki/Wikidata:WikiProject_LGBT/gender.

Wikimedia Foundation. n.d. "Our Values." Accessed July 1, 2021.
<https://wikimediafoundation.org/about/values/>.

Appendix: 50+ Things You Can Do to Get Started with Wikidata in Your Archives²

This list identifies over 50 simple steps you can take to get started working with Wikidata in your Archives. While the simpler steps are near the beginning, one can take these steps in any order. *50 Things* was created Kelli Babcock, Regine Heberlein, Anna Björnsson McCormick, Elizabeth Russey Roke, Greta Kuriger Suiter, and Ruth Kitchin Tillman. This list builds on the "[Collections As Data: 50 Things You Can Do](#)" list item number 47: *Explore what it would take for your organization to contribute subject data to Wikidata, drawing on a local collection and then incorporating the Wikidata links into your local discovery environment.*

We have published this list as [a living document](#) in the Archives Linked Data Interest Group WikiProject space and encourage others to contribute.

1. **Getting started? First...** Don't be afraid to make mistakes!
2. Create a Wikidata account at [Special:CreateAccount](#)
3. Not sure how to create a Wikidata account? Start learning more about Wikidata from the free [WikiEdu training documentation](#)
4. Consider paying for a class with [Wiki Education](#) for facilitated training and support
5. Consider watching [Wikidata:WikiProject LD4 Wikidata Affinity Group](#) meeting [recordings](#) (stored in Google Drive as of July 2021)
6. Don't have enough time to read through all of that training? Check out Heather Dean's [explanation slides for introducing Wikidata to archivists](#)
7. Check out the [Help:Contents](#) to learn more about how to get started in Wikidata
8. Learn by doing!
9. Pick a Wikidata item of your choice—or a random item from the menu on [Wikidata:Main Page](#)—and evaluate it
10. Pick a Wikidata item of your choice—or a random item from the menu on [Wikidata:Main Page](#)—and add a property (see [Wikidata:List of properties](#))
11. Check out the [property explorer](#)
12. Not sure what other items to link to your item? Check the "What links here" on the left and fill in the gaps.
13. Use the [Wikidata sandbox](#) to test out your changes
14. Don't worry about conflicting statements; allow them to stand
15. Play [Terminator](#) to practice editing items
16. Learn about the [Wikidata graph builder](#)
17. Join a WikiProject, such as [Wikidata:WikiProject Archives Linked Data Interest Group](#) by following
18. Not sure how to use Wikitext markdown? Neither were we at first. Check out [Help:Wikitext](#) for an introduction

² "[Wikidata:WikiProject Archives Linked Data Interest Group/50 Things](#)", republished under [CC BY-SA 3.0](#). Introductory text was modified slightly by original authors.

19. **Start to describe your repository and archival materials in Wikidata**, but first: go to your Wikidata preferences, select the subpage for gadgets, and turn on gadgets for:
 1. DuplicateReferences
 2. EasyQuery
 3. currentDate
 4. explore edit tools at [Wikidata:Tools/Edit items](#)
20. Before you begin adding "archives at" statements in Wikidata, you will need a Wikidata item for your repository. Look up the Wikidata item for your archives/repository in the "Search Wikidata" top right search box. Compare it to a thoroughly-described repository like [Seeley G. Mudd Manuscript Library \(Q7445830\)](#). Look at other archives repository Wikidata item examples. Note what you like about different practices across repositories
21. If a Wikidata item for your repository exists, consider how you can add properties to the item based on documentation like [Wikidata:WikiProject Archives Linked Data Interest Group/Describing Repositories](#)
22. If no Wikidata item exists for your repository, review [Wikidata:WikiProject Archives Linked Data Interest Group/Describing Repositories](#) and make a basic stub record for your repository. Remember that your item can be as simple or full as you would like. [Modelling GLAM institutions on Wikidata](#) can also help you decide how to describe your Wikidata repository item
23. Look up your repository item in [the Finding GLAMs Tool](#)
24. Find other repositories near you using the Finding GLAMs Tool and consider reaching out to them. You can also look at other repositories in your country who are using Wikidata "archives at" statements with [this](#) SPARQL query ("Canada" is in the canned query - you can change [country \(P17\)](#) in the query depending on where you are)
25. Now pick your favorite fonds or collection in your repository. Search for the person/organization who created the collection in Wikidata. If the person/organization has an existing Wikidata item, add the [archives at \(P485\)](#) property to the item and link to your Wikidata repository item. It is also useful to add a [described at URL \(P973\)](#) link to the finding aid (even better if it has a DOI) and a [owned by \(P127\)](#) for the collection's identifier. If the person doesn't exist...
26. ...Create a Wikidata item for the person/organization, then add an [archives at \(P485\)](#) Wikidata property to the item with a link to your repository Wikidata item. See [Wikidata:WikiProject Archives Linked Data Interest Group/Description Recommendations](#) for guidance on adding people and organizations items to Wikidata
27. Become familiar with your local practices around authority records by conducting a local agent record survey
28. Evaluate and update your local agent record documentation
29. Try this [reconciliation tool](#) to improve your authority records
30. Create a spreadsheet of people and organizations whose papers are held in your repository. Maintain this list, including recording how you created it:
 1. If your institution uses [AtoM](#), you can add authority records [to the clipboard and export as csv](#), or use [this MySQL query to export](#)
 2. If your institution uses [ArchivesSpace](#), you can select Agent records and export as csv
31. Update your arrangement and description workflows to include adding a P485 property to the entity's Wikidata item

32. Know that you do not need to create the "perfect" Wikidata item for each person. Take a "minimum viable product" approach as you are learning
33. Create or adapt data models for persons, organizations
34. Know as you are doing this work that the data models and best practices can be defined by you, with guidance from the Wikidata - and Wikidata + archives communities - to meet Wikidata community guidelines and also meet the context of your particular project
35. Add robust biographical or historical notes to collection descriptions; this information can be repurposed in Wikidata
36. Evaluate and update biographical and historical note documentation
37. Add Wikidata identifiers to the bibliographic/authority records for collection creators in your local catalog or archival management system
38. **Start sharing your knowledge about Wikidata.** Talk to others in your organization - Wikipedia enthusiasts, catalogers, IT staff. Find others interested in linked data
39. Open a Slack channel dedicated to discussing Wikidata within your organization
40. Create a local affinity group in your organization - set up monthly/quarterly meetings, encourage group learning and sharing of expertise
41. Host a Wikidata edit-a-thon at your local institution to add entities related to your institution to Wikidata
42. Offer some informal Show & Tell within your organization or group
43. Find out if there is a Wikimedia group or project that relates to topics you are interested in or local geographic based groups. For example: Women in Red, Black Lunch Table, AfroCrowd, Art+Feminism, Los Angeles group, Colorado group, etc. Find more groups [here](#)
44. **Take it further and start querying and analyzing Wikidata!** Learn about the [Wikidata SPARQL Query Service](#) and play with the sample queries
45. Take a sample query, make a simple edit, and run it in the [web interface of the query service](#)
 1. The sample queries below can have their [archives at \(P485\)](#) changed to any repository. The canned queries are querying against archives at for [University of Toronto Archives & Records Management Services \(Q64825166\)](#) and are shared from [utl awong](#):
 1. Querying [University of Toronto Archives & Records Management Services \(Q64825166\)](#) "archives at" for instance of: <https://w.wiki/nze>
 2. Querying [University of Toronto Archives & Records Management Services \(Q64825166\)](#) "archives at" for sex/gender: <https://w.wiki/nzf>
 3. Querying [University of Toronto Archives & Records Management Services \(Q64825166\)](#) "archives at" for occupation: <https://w.wiki/nzj>
46. Take a [SPARQL tutorial](#)
47. **Continue learning!** Learn about QuickStatements for bulk uploading - [10 minute tutorial](#)
48. Download OpenRefine and get started using [Thomas Padilla's walkthrough](#) (data included)
49. Watch [Owen Stephens' videos on reconciling to Wikidata in OpenRefine](#)
50. Learn about QuickStatements: [Help:QuickStatements](#)
51. Learn about lists with [Listeria](#)
52. View Wikidata in table format using [Tabernacle](#)
53. Try a form to create a Wikidata item with [Cradle](#)

Inviting and Honoring User-contributed Content

Katherine Crowe, Katrina Fenlon, Hannah Frisch, Diana Marsh, and
Victoria Van Hyning

Abstract: Galleries, libraries, archives, and museums (GLAMs) and users deploy and engage with many emergent technologies, platforms, projects, and approaches to gather *user-contributed content* (UCC) related to GLAM collections, such as crowdsourcing, linked data, and tools and platforms for digital scholarship. However, actually integrating UCC with long-term preservation, search, and discovery systems lags behind. Though there are many challenges to integrating UCC, the demonstrable benefits of integration are significant and can promote broader collections use and impact.

Introduction

This contribution considers how various cultural institutions solicit and integrate *user-contributed content* (UCC) to improve discovery, access, use, and accessibility of their collections, while deepening engagement with a wide variety of stakeholder communities. We review the state of professional practice, exciting opportunities, and considerable challenges of leveraging UCC, including issues of quality control, technical interoperability, user incentivization, digital preservation, and sociotechnical workflows. We focus on case studies including (a) crowdsourcing initiatives; (b) community collaborations and collections, focused on Indigenous collections; and (c) generative user practices particularly in the digital humanities, and obstacles to sharing researcher-generated data. Our objectives are to shed light on existing practices of UCC integration, identify common obstacles and offer suggestions for how to tackle them, and highlight the ethical and moral issues of not pushing this work forward.

Opportunities and challenges of user-contributed content

In an important study of how GLAM institutions solicit and integrate what she calls “social metadata” and “public-contributed content” (and which we refer to collectively here as “user-contributed content” or UCC), Chern Li Liew (2016) finds that while many institutions solicit data from users—such as tags, notes, images, personal memories—few then integrate these materials into their core knowledge infrastructure. In her study of Swedish GLAMs and crowdsourcing, Ina-Maria Jansson (2017) finds similar gaps in integration of the data resulting from public engagement with GLAM resources, tracing the disconnect partly to inflexibility in GLAM metadata systems (also found by Blaser 2014), and in part to anxieties about how or whether to integrate unverifiable or subjective information, such as users’ personal memories, into authoritative records (2017).

While questions of how and whether to integrate data persist, many institutions are continuing or beginning to actively solicit UCC through diverse programs and measures including crowdsourcing, comment and metadata fields in catalogs, paper forms in reading rooms, user-photography and scanning labs, and Wiki edit-a-thons. Potential UCC includes folksonomic tags; transcriptions, data, and other content from external crowdsourcing platforms; community stories, contextual information, and primary sources gathered from communities through digitization events; etc. Many GLAM practitioners recognize UCC as an opportunity for engagement, but on the whole the field has yet to realize the full potential of UCC for organizations or users.

UCC is contextualized by a wider movement in the literature and practice of archives, toward deeper and more varied forms of community engagement, e.g., through models of participatory archiving (e.g. Huvila 2008; Shilton and Srinivasan 2007; Iacovino 2015; Rolan 2017), postcustodial archival practices (Bastian 2002; Cunningham 2011), and shared stewardship between communities and institutions (e.g. Smithsonian Center for Folklife and Cultural Heritage 2019; Powell 2016). The importance of community involvement and co-creation of archival collections is well documented by scholars in the field (Evans 2007; Roued-Cunliffe and Copeland 2017; Somerville and EchoHawk 2011). Including UCC deepens community engagement and empowers communities by providing space for the public's expertise and knowledge to add value directly to collections (Christen 2011; Ridge 2014; Alemu 2018; Roued-Cunliffe and Copeland 2017; Somerville and EchoHawk 2011; Stevens, Flinn, and Shepherd 2010).

In addition to dedicated community engagement initiatives, everyday archival user practices generate content with potential value, both for other users and for institutional collections. Researchers using physical and digital archival collections create a variety of content in the course of their use, documentation, annotation, and interpretation of collections. For example, reading room photography has increased dramatically as scholarly, family, Tribal, and local researchers gain access to inexpensive, high-capacity cameras and institutions lower barriers to imaging. As users collect and annotate images during their research, they generate byproducts such as digital images of analog sources, additional metadata via annotations, etc., which institutions might leverage to increase the value of and access to collections.

Though there are countless forms and many benefits of UCC, there are numerous obstacles to its widespread implementation in institutional collections. Studies have described reluctance to rescind control of processing and describing materials. Alemu (2018) describes the reluctance of professional archivists to relinquish their institutional gatekeeping responsibilities. Collaborations between institutions and communities tend to founder on a lack of shared understanding of mutual expertise (Flinn 2011). Concerns about data quality impede admittance of UCC into archival description practices (van Hooland, Méndez, and Boydens 2011; Liew 2016). Frequently, content management systems (CMS) or data models undergirding institutional systems lack the capacity to admit new data types entailed in UCC, and normalizing, cleaning, and parsing UCC is a notoriously difficult and resource-intensive task (Blaser 2014; Bullard 2019, Patterson III 2012, Alemneh and Hastings 2008, Buckland et. al 1999; Frommholz et al. 2014; Jansson 2017; Tonkin and Tourte 2016; Matienzo and Rudersdorf 2014; Thompson 2017).

The obstacles are considerable, but the benefits of UCC are more considerable still, and growing. In this paper, we consider cases of UCC in different contexts, including crowdsourcing initiatives, community collaborations and collections (focused on Indigenous collections), and the potential of generative,

everyday practices of archival users to contribute meaningful content to collections. Through these cases, we shed light on existing practices related to UCC, identify common obstacles to integrating UCC into institutional collections, and suggest the ethical and moral imperatives of increasing UCC in institutional collections.

Crowdsourcing UCC

GLAMs have a long tradition of supporting on-site volunteer programs that engage people in processing, cataloging, description, deaccessioning, and digitization as well as collection interpretation through docent work, education, and exhibit curation. These valuable contributions by volunteers increase the accessibility and usability of collections and expand historical context and educational opportunities. Likewise, many academic institutions, dictionaries, and encyclopedias have engaged the public in research and the creation of authoritative reference sources for over a century. The *Oxford English Dictionary* is an oft-cited example (Lanxon 2011; Barber 2017). Beginning his 70-year project in the late 1800s, editor James Murray solicited examples of word usage from the public and was sent slips of paper by thousands of people containing citations, definitions, contexts, and dates for early or unusual usages (Mugglestone 2005). Many biographical dictionaries, such as the *American Dictionary of National Biography*, rely on a large, distributed network of experts, some with professional or academic qualifications, and some with deep knowledge acquired through self-study or lived experience. Wikipedia is simply a permutation of, rather than a radical departure from, this long tradition of distributed, collaborative knowledge creation.

Online crowdsourcing, also known as virtual volunteering, is an extension of the practice of on-site volunteering that allows for more flexible and geographically distributed collaboration among GLAM institutions, researchers, educators, and a global volunteer base. Over the last two decades, crowdsourcing platforms and tools have been adopted or created by a large number of GLAMs as they seek to enrich metadata about their collections while building deeper relationships with communities (Blaser 2014; Ridge 2014; Owens 2014; Hedges and Dunn 2018). Online crowdsourcing can include many activities, including transcribing, tagging, commenting, translating, correcting OCR'd records, and Wikipedia edit-a-thons. The National Archives and Records Administration's (NARA) Citizen Archivist, the Library of Congress' By the People, and the Smithsonian Transcription Center are national institutions that have created new platforms to solicit transcriptions and tags of national collections, while many smaller organizations in the US and Europe in particular have adopted platforms such as Zooniverse, PyBossa, and From the Page, which engage large volunteer bases, as well as providing DIY tools for GLAM, academic, and other users to build their own crowdsourcing projects.

Mia Ridge's definition of cultural heritage crowdsourcing encapsulates the mutual benefits inherent in these endeavors (whether on-site or online), which "ask the public to undertake tasks that cannot be done automatically, in an environment where the activities, goals (or both) provide inherent rewards for participation, and where their participation contributes to a shared, significant goal or research area" (2014, 2). Writing about *Old Weather*, an early and very successful citizen science project featuring historic ships' logs, Lucinda Blaser (2014) discussed the value of crowdsourcing in advancing scientific endeavors while also enhancing cultural heritage content. Yet she also sounds a note of caution and is one of the few GLAM practitioners to have done so, to date: "Of course crowdsourcing and citizen science cannot be seen as a solution to all problems that cultural institutions face with their collections. These projects present their own set of problems, such as the best way to incorporate the generated

data back into the institution's systems" (24). Thus while the benefits of crowdsourcing are well-attested and the motivations of many cultural heritage practitioners to enhance their collections and engage a broad public are genuinely admirable, the challenges of integrating UCC into content management systems persist. If these challenges continue to go unaddressed, this would constitute a significant failure on the part of GLAMs to uphold the often explicit promise they make to volunteers: that the data they create will help enhance access to collections and further research. Numerous studies demonstrate that volunteers contribute altruistically, hoping to help researchers of today and the future (Jennett et al. 2016; Cox et al. 2018).

Virtual crowdsourcing is a relatively new practice in cultural heritage (Holley 2010), but it is mature enough and widespread enough for us to address the prevalent disconnect between the public solicitation of volunteer time and effort by GLAMs, and widespread failure to meaningfully integrate the resulting data (Liew 2016; Jansson 2017). It is likely that many socio-technical solutions implemented to fulfill the promise of cultural heritage crowdsourcing would also significantly enhance GLAMs' ability to ingest other forms of UCC in order to enhance collections access and usability.

Potential models for integrating UCC

This section will discuss crowdsourcing and UCC initiatives at several large federal institutions and the Folger Shakespeare Library, with which the authors are particularly experienced. Good practice exists elsewhere, and these examples are purely indicative of different models rather than the only ways of integrating UCC.

NARA exemplifies the fusion of on-site crowdsourcing, virtual crowdsourcing, and the successful integration of UCC to promote access and discovery. In addition to hosting on-site and virtual Wikipedia edit-a-thons, the on-site [Innovation Hub](#) is a space where members of the public can scan NARA records for their own use, which are then integrated into NARA's catalog. Once in the catalog, the materials are available for volunteer transcribing and tagging on NARA's [Citizen Archivist](#). Unlike many stand-alone crowdsourcing projects listed above, Citizen Archivist is coupled with the catalog, which means that UCC immediately appears alongside the content it enhances and can be understood as part of the authoritative record.

The Smithsonian (Orli and Bird 2016) and Library of Congress (Van Hyning 2020) have had to create new pathways to publish crowdsourced transcriptions and tags from their stand-alone sites alongside core institutional web properties. At the Library of Congress, transcriptions are created and reviewed by volunteers on [crowd.loc.gov](#). When a 'Campaign' or collection of documents is complete, subject specialists spot-check 5% of the Campaign, looking for incomplete transcriptions or errors that will hamper search, such as mistranscribed proper nouns. Once data goes through QC, they are packaged as datasets for loc.gov, as well as linked to individual item- and page-level records and images. An attribution is included in each searchable, downloadable .txt file, and an overlay appears over the transcription viewer on the Library of Congress website (e.g. [in the Clara Barton Papers](#)) stating 'Transcribed and reviewed by volunteers participating in the By the People project at crowd.loc.gov' (Van Hyning 2020).

At the time of writing, Past Perfect, like many commercial CMS products, only allows manual copy and paste into free text metadata fields, but there are no bulk import options to import a new metadata

field into existing records. For large scale UCC projects, this can be a helpful quality control mechanism, but for others it is a significant barrier to the inclusion of UCC.

NARA hosts a separate online open reference and discussion platform called [History Hub](#), for answering researchers' and volunteers' questions. Open reference and discussion platforms are significant arenas for what Holley (2010) terms "social engagement." Distinct from UCC or crowdsourcing, this related form of interaction between the public and institutions allows volunteers and researchers to aid one another, as well as seek institutional guidance.

Digitization and curatorial specialists at the Folger Shakespeare Library in Washington DC have been inspired by NARA's Innovation Hub and are taking it a step further. Closed for renovations from 2020 to 2023, the Folger is piloting a project to ingest several sub-preservation grade images sources into their searchable image repository: reading room photography often shot on smart phones by visiting researchers for their future reference, and reference photography created by staff to share with remote users via email. The photos may not be preservation quality or conform to best practice, but for sighted users they are a huge boon to research; they can in turn be used in one of the Folger's current crowdsourced transcription projects on From the Page (<https://fromthepage.com/folger>), through which volunteers and staff create searchable and accessible transcriptions in the original spelling of the documents as well as modernized spellings. The Folger unites these heterogeneous resources into their web properties to create full-text finding aids that fundamentally reimagine the role and function of traditional finding aids (Wolfe and Van Hyning, forthcoming). The latter have been helpfully described and roundly critiqued by Wiedeman (2019), Daniels and Yakel (2010), Daines and Nimer (2011; 2008), Scheir (2006), and others.

To integrate UCC into collections is to fundamentally expand and shift the meanings of institutional authority. Democratizing collection processing, enhancement, and description work empowers users to access and engage with collections in ways that may promote a sense of shared ownership and responsibility. In the next section we consider emerging modalities of stewardship and repatriation of Indigenous cultural heritage.

Indigenous collections and community collaborations

Broader movements in community archiving, reparative description, and postcustodial archival stewardship are supporting the incorporation of community-created content with institutional collections within community repositories. Shared stewardship (Smithsonian Center for Folklife and Cultural Heritage 2019), ethical collaboration (Genovese 2016; O'Neal 2019), digital return (Barwick et al. 2019; Bell et al. 2013), knowledge sharing (Carpenter 2019; Powell 2016), "repatriation" (Christen 2011), and wider adoption of the *Protocols for Native American Archival Materials* (First Archivists' Circle 2007) are making possible new partnerships between institutions and communities, and generating models in which communities own, steward, or curate institutionally-held collections in their own digital or physical spaces.

Policies facilitating the return or sharing of institutional collections, alongside long-standing work on the part of community researchers to gain access to copies of institutional collections, are allowing huge digital and physical archives to be built in communities. These wholly community-curated repositories offer greater flexibility than institutional repositories (even those with forward-thinking

policies), such as the ability to enhance, annotate, translate, or protect collections based on community-generated knowledge, and the ability to bring together researcher images with institutional surrogates as well as community-created digital collections such as oral history recordings, locally digitized images, or local film footage (see, for instance, Christen et al. 2017; Dallwitz et al. 2019).

Creating those on-site repositories is essential for community uses of archives. Existing research has shown that many communities for whom archival documents are essential to cultural revitalization and community power are least able to access them (Marsh 2018). Even in the pre-pandemic world, many community-based researchers and non-academics found physical access to archival reading rooms cost-prohibitive, if not culturally impenetrable (Buchanan et al. 2021). Major GLAM repositories, usually located in cosmopolitan centers, are often far from the communities whose knowledge and histories their collections represent (See First Archivists' Circle 2007, "Building Relationships of Mutual Respect").

To interpret historical Indigenous language materials held in archives, many Indigenous community research trips, for instance, must include groups of researchers—such as linguists and fluent speakers. The latter are usually elders, whose time and knowledge are a community's most precious resource. (The Recovering Voices program at the Smithsonian budgeted \$10,000 for its competitive awards to fly these groups to Washington, DC to undertake collections research for a single week.)

It may be some time until such travel for research endeavors is safe for elders, many of whom have already been lost to the COVID-19 pandemic. Now, the GLAM field has an additional ethical responsibility to provide digital access to collections so that research can be done at home in Native and Indigenous communities.

Community researchers do not need access to preservation-quality digital images; they need access to the information held in collections. That information may well already have been documented by other researchers to those collections but is dispersed among individuals across the globe.

Many longstanding efforts in crowdsourcing of images and metadata for use in community and digital archives have shown the utility of even low-resolution imagery for reuse outside of institutional contexts. Many Native and Indigenous community centers such as the [Deyohahá:ge Indigenous Knowledge Centre](#) at Six Nations Polytechnic, use reading room photographs from archives around the world (alongside formally acquired institutional surrogates) in their digital archives; such archives use their own community-based CMS and ontology in which to organize and describe these collections for ease of community use.

That effort is part of a broad movement in the creation of Indigenous archives and databases that crowdsource community knowledge to interpret, organize, and responsibly steward collections. Many of these projects involve a specialized and community-curated CMS, such as [Mukurtu CMS](#), the [Great Lakes Research Alliance for the Study of Aboriginal Arts and Culture \(GRASAC\) Knowledge Sharing System](#), and the [Ara Irititja Keeping Culture Knowledge Management System](#). These systems are designed with the flexibility to incorporate community members' annotations and metadata enhancements; to place collections under cultural restrictions using traditional knowledge labels or alternative copyright (see

<https://localcontexts.org>; Anderson and Christen 2013), and to steward community images or recordings alongside institutional surrogates.

Similarly, intentioned archives are being created to hold, annotate, and translate Indigenous language materials for mobilization in language revitalization, such as the [Indigenous Languages Digital Archive \(ILDA\)](#), part of the Myaamia Center at Miami University, and the University of Alaska Fairbanks' [Alaska Native Language Archive](#). Linguistic databases working to aggregate Indigenous linguistic archives, such as the [Archives of Indigenous Languages of North America](#) (AILLA). As in other digital humanities realms, scholarly projects such as the J. P. Harrington Database Project at the Native American Language Center, Department of Native American Studies, University of California, Davis, work to transcribe Indigenous cultural and language materials held in faraway institutional repositories (Macri et al. 2009).

Mukurtu has recently been funded to pilot a cloud-based platform, Mukurtu Shared. That platform, if successful, will ingest repository images, allow community annotations and interpretations of them, and in turn share that metadata back with repositories to be incorporated into institutional CMSs. But accomplishing that goal remains challenging. Most community-based platforms are not interoperable with major institutional CMSs, whether these are bespoke creations or off-the-shelf tools such as ArchivesSpace. Instead, much of this translational work of bringing institutional content into a community CMS, or community content back into institutional systems, will require data file exports and imports, completed manually by community and institutional archivists (on these broad challenges see Matienzo and Rudersdorf 2014). Moreover, institutional systems like ArchivesSpace often do not have enough built-in flexibility to account for community annotations, additional metadata fields, new thesauri, or other content.

Meanwhile, most institutional repositories do create their own internal “photocopy scans” for use in reference requests. Yet, while those surrogates might be shared with communities—and incorporated into systems like Mukurtu or other community repositories—at the holding institution (as described at Folger and NARA above), they are often kept completely separate from digital collections, and not available publicly for reuse and wider access. In the next section, we consider broader trends in archival users’ content, particularly in digital humanities and scholarly work.

Generative user practices and obstacles to sharing

The everyday research and publishing practices of scholars relying on archives tend to generate content ripe for leveraging as UCC. The research processes of scholars and other archival users are complex: they wend toward shifting intellectual objectives, governed in part by disciplinary and methodological imperatives and norms, and in part by individual idiosyncrasies (Borgman 2015; Palmer et al. 2009). For several decades, empirical information practices research and user studies research in library and information science and archives have sought to characterize these processes to guide research support and systems development. From the study of these processes, information practices research has distilled various frameworks of common activities, basic building blocks of scholarly work, sometimes called *primitives* (Unsworth 2000). These include activities such as “browsing”, “reading” and “re-reading”, “networking”, “collaborating”, “writing”, etc. (Ellis 1993; Unsworth 2000; Palmer et al. 2009; Benardou 2010; Blanke and Hedges 2013; Vilar 2015).

The most generative scholarly activities, from the perspective of UCC, are *note-taking* and *collecting*: scholars create annotations throughout their research processes, and simultaneously create personal collections of materials to support their work. *Note-taking* or annotation pervades most research processes, from the outset of searching to the formative stages of writing—producing marginal notes, documents, sketches, maps, outlines, and chronologies of varying levels of structure and formality (Palmer et al. 2009). While most note-taking and annotation activities are private and intended to support individual intellectual tasks, a growing number of systems support sharing and collaboration in note-taking and annotation, both within research teams and across wider research networks (e.g., collaborative note-taking features in most citation management systems, Hypothes.is web annotation, Roam, etc.). For example, historians curate structured datasets in spreadsheets by reading and interpreting archival documents; literary scholars select and transcribe quotations and contextual information; and anthropologists select and document observations of field sites. In addition, numerous studies have documented the collecting practices of scholars in various disciplines (Brockman et al. 2001; Palmer 2004; Palmer et al. 2009; Trace and Karadkar, 2016; Cooper and Rieger 2018; etc.). Scholarly collections may be large, and organized and documented to varying degrees. They are increasingly digital. Across disciplines, collections may include original or duplicated representations of evidence: datasets, field notes, recorded interviews, visual materials etc. In disciplines relying on archival evidence, scholars' personal collections usually include copies or photographs of archival resources from various institutional collections, along with relevant secondary sources.

Studies have long suggested that scholars' personal collections hold potential value for other researchers, yet sharing personal collections remains rare in practice (Spanner 2001; Trace and Karadkar 2016; Cooper and Rieger 2018). There are numerous obstacles to sharing collections of primary- and secondary-source materials gathered from archival collections (all closely related to the well-documented obstacles to sharing research data more generally). Scholarly credit and promotion systems do not reward data sharing, and users may lack time or resources to document or convert sources adequately for use by others. Their sources may offer a competitive advantage in research, scholars may understand sources as their intellectual property, or may be unable to assess privacy and IP-related considerations. They may lack access to appropriate platforms for sharing, and many scholars simply remain unpersuaded of the benefits or value of sharing broadly (Borgman 2010; Poole 2013; Cooper and Rieger 2018). When scholars working with archival materials do share their collections, they tend to do so within their personal networks, often through direct contact with colleagues, and very rarely with original collecting institutions. Systems to support data- and source-sharing among humanities scholars are increasing, with the emergence of models for data citation, systems for according credit for curatorial work, and gradual changes to tenure and promotion processes to reward different modes of labor in digital scholarship. Yet, incentives for data- and collection-sharing among scholars are far from established; sharing remains disincentivized by technical, social, and institutional factors in the prevailing paradigm of humanities publishing.

Additional obstacles to sharing arise from the ethos and norms of professional archival practice. The Society of American Archivists' *Code of Ethics for Archivists* asserts that archivists should maintain the confidentiality of users' research. While necessary in many cases, this mandate undermines the role archivists might play in serving as hubs for collaborative research or facilitating scholarly communication and resource-sharing (and indeed, most reference archivists find ways to work around this principle ethically, such as by asking researchers if they wish to be connected).

Digital humanities projects and publications amplify opportunities for connections between institutional and community collections. Wide-ranging efforts—from highly detailed digital editions, to interactive maps, to large-scale databases of structured, historical data—tend to produce or collect unique or original evidence, often of groups and histories that are underrepresented in mainstream cultural institutions. Like community archives, these collections often live independently of libraries and archives, but are related to institutional collections to varying degrees. Digital scholarly editions represent complex, richly encoded versions of manuscripts and other original sources held in archives. Digital humanities databases aggregate computationally amenable representations of information from scattered historical sources. Digital archives produced by scholars include high-quality, digitized versions of primary sources with a wealth of added contextual information and metadata. Furthermore, there are usually additional, value-added materials associated with digital humanities collections that could dramatically enrich the context and documentation attached to original collections, but which never make it back to the originating institutions. For example, the same digital humanities archive relying on an institutional API might have created rich, multifaceted Text Encoding Initiative (TEI) encodings of its sources, or rich Geographic Information Systems (GIS) data pertaining to its sources, which materials and added data the institution has no way of including in its collections. Such resources may prove immensely valuable to other researchers. Sometimes, these projects include actionable links to the collections they derive from or relate to. For example, a digital archive may provide hyperlinked citations to original sources in an archival collection, or may even call on a cultural institution's IIIF API to provide access to digital images in a thematic collection without removing the images from their original institutional context. Such actionable or explicit connections between digital scholarship and cultural collections remain rare in digital humanities practice, however.

As cultural institutions rethink their roles in and responsibilities toward communities more generally, they have the opportunity to rethink how their collections are related to the wider world of cultural knowledge, as manifested in social networks, digital scholarship, publications, data repositories, etc. Could institutions forge actionable, useful links between their own archival collections and external, scholar-generated datasets or scholarly editions derived from those collections, for example? Could institutions play a meaningful role in connecting scholars working with the same resources to reduce duplicated efforts? What if more GLAMs attempted to unite researchers' reading room images, notes, transcriptions, and datasets derived from collection materials, a great deal of which are currently hosted and gently curated by researchers on their personal devices and clouds, with the institutional collection? The incentives for users to share their resources are arguably weak at present (Borgman 2015), but a mix of encouraging and modeling the desired behavior, offering incentives such as DOIs for data, periods of embargo (for users' notes, transcriptions, or other resulting data), fellowships, seminar talks/webinars, and opportunities to deposit and link resulting research publications that feature the institutions' holdings may help to turn the tide.

Technologies to support widespread sharing of and interconnection among research materials exist—some emergent, others well established—including social media (and attendant web- and social-media archiving practices), collaborative software, linked data, annotation platforms, etc. Despite their concerns about corporate ownership of platforms, scholars use social media to share digital archival sources with meaningful metadata and annotations (Cooper and Rieger 2018). While some cultural institutions are making forays into social-media and Web archiving, this practice remains rare; and it is even more rare that social media content would be used to augment or enrich metadata in extant

collections. Open source software for managing research materials—ranging from reference management systems like Zotero to image- and metadata-management tools like Tropy—increasingly support public sharing and collaborative work, opening the possibility of publicly accessible masses of research data and digital materials. Increasingly prevalent linked data standards and tools can support virtual gathering, by forging meaningful, actionable, useful connections among scattered collections of content, including between user-contributed and institutional collections. Institutional experiments with this are underway. Web-based annotation, tagging, and note-taking tools, like Hypothes.is, Roam, and TagTeam, many of which are based on linked data or semantic technologies, have proven value for supporting scholarly communication; and some cultural institutions have made forays into supporting user-contributed, shared annotations within digital collections (e.g., Oldman and Tanase, 2018). The growth of data journals, open-access digital publishing platforms, and shared infrastructures for digital scholarship in the humanities suggest growing scholarly interest in sharing research materials.

These technologies evince a widening array of opportunities for gathering UCC, but do not address certain basic challenges to integrating UCC meaningfully with institutional collections, including resource constraints (posed against an already burgeoning mass of unprocessed collections); a lack of workflows for integrating UCC into archival collections; and the overriding the need for archival practices and ethos to grow in capacity and flexibility to accommodate community expertise and contributions. Nonetheless, there are increasing models of success. In addition to the case studies offered here, the examples of Europeana, RunCoCo, and the Digital Repository of Ireland’s institutional scanning programs may offer useful models. These groups invite members of the public to scanning open days where their personal archival documents can be scanned and then integrated into institutional data structures, where they undergo description, data normalization and aggregation. Efforts such as Digibird, which provides workflows, APIs, and methods for data integration (Dijkshoorn et al. 2017), may likewise offer valuable insights for practitioners seeking models for data integration.

As a growing wealth of digital scholarship and scholars’ personal collections remain mostly disconnected from institutional collections, a certain amount of value is wasted. In particular, the disconnect between scholarly materials and the institutions in which materials originated represents a lost opportunity, both for the impact of scholarship and for the value of original collections. Work in the domains of data curation and library research services over the past decade has grown to support data- and collection-sharing within scholarly communities and with broader publics. A wealth of literature and guidance on sharing to support scholarly reuse, particularly through data repositories, has emerged in the past decade, but leaves a basic question unanswered: What is the role of cultural institutions in connecting scholarly collections to institutional collections? How can they help reclaim the value of scholarly efforts around their own archival sources? How can more informal modes of sharing, which are low-overhead for scholars and institutions, improve primary collections and deepen scholarly engagement with cultural institutions? For example, institutions might create and advertise opportunities for researchers to donate or deposit their reading room photography, transcriptions or even copies of any final publications that draw deeply on the institution’s materials. Attribution and acknowledgement in the metadata associated with the primary sources, formal recommendations for subsequent users in acknowledging the photographer, small researcher stipends and a variety of other mechanisms could be established to support a shift in the current landscape.

Conclusion

As we write in the summer of 2021, the question of whether and how the cultural heritage community might gather and share UCC is more pressing than ever. The almost universal closure of GLAMs around the world due to the COVID-19 pandemic now means that digitized collections are often the only form of viable access for researchers, students, and a variety of users from all walks of life. At other times, routine closures, renovations, reduced operations or the relocation of materials and sites can curtail access. War and other disasters can cause significant loss, dislocation or complete destruction of collections, as well as sometimes gravely harming or killing the people who use them, know most about them or who have the deepest cultural connection with them.

The devastating fires at the National Museum of Brazil in 2018 and the University of Cape Town in 2021, which resulted in the loss of over 300 million artefacts and archival materials between them, provide extreme recent examples of such loss. Both events also provide examples of post-disaster crowdsourcing efforts (Boyle 2018; SAA 2021)—a variety of professional, academic, student-led, and community-led efforts to gather UCC such as photographs, films, 3D images of objects, as well as transcriptions and descriptions of lost texts, and other knowledge (for example researchers' notes, community oral histories, and more). The public response to these calls is heartwarming to be sure, but a desire on the part of users to share knowledge and help institutions is not limited to the aftermath of disaster. As countless crowdsourcing and community archiving efforts around the world demonstrate, goodwill is there in abundance, stemming from people's desire to contribute to knowledge and research, and help make materials accessible for people with a range of disabilities.

The creation and management of cultural knowledge happens in many domains outside of cultural institutions, and the need and opportunities to unite community and institutional knowledge are growing. The case for this union strengthens as the precarity of cultural collections outside of cultural institutions (such as digital humanities scholarship and community archives) becomes clearer, and as the benefits to institutional collections are made visible by a growing number of initiatives.

The sociotechnical processes and workflows that govern how collections are created, managed, connected, and made accessible pose the most immediate challenges and research opportunities for the advancement of UCC: What technologies and human-centered processes or workflows can best facilitate collaboration between users and keepers of collections at your institution and in your discipline? How do these processes or workflows become viable and then made sustainable? What are the roles involved and who plays them? Where are the gaps?

Given the conducive shift in archival and broader GLAM ethos and practices—toward community-engaged work—the time is ripe for empirically researched, systematic approaches to incorporating UCC into institutional collections, and a broader commitment, particularly by those already engaged in soliciting UCC, to find and share solutions to preserving these vital forms of knowledge.

References

- Alemu, Getaneh. 2018. "Metadata Enrichment for Digital Heritage: Users As Co-Creators." *International Information & Library Review* 50 (2): 142–56. <https://doi.org/10.1080/10572317.2018.1449426>.
- Anderson, Jane, and Kimberly Christen. 2013. "'Chuck a Copyright on It': Dilemmas of Digital Return and the Possibilities for Traditional Knowledge Licenses and Labels." *Museum Anthropology Review* 7 (1-2): 105-26.
- Barber, Samuel T. 2018. "The Zooniverse Is Expanding: Crowdsourced Solutions to the Hidden Collections Problem and the Rise of the Revolutionary Cataloging Interface." *Journal of Library Metadata* 18 (2): 85–111. <https://doi.org/10.1080/19386389.2018.1489449>.
- Barwick, Linda, Jennifer Green, and Petronella Vaarzon-Morel. 2019. *Archival Returns: Central Australia and Beyond*. Sydney: Sydney University Press.
- Bastian, Jeannette A. 2002. "Taking Custody, Giving Access: A Postcustodial Role for a New Century." *Archivaria* 53: 76-93.
- Bell, Joshua A, Kimberly Christen, and Mark Turin. 2013. "Introduction: After the Return." *Museum Anthropology Review* 7 (1-2): 1-21.
- Benardou, Agiatis, Panos Constantopoulos, Costis Dallas, and Dimitris Gavrilis. 2010. "Understanding the Information Requirements of Arts and Humanities Scholarship." *International Journal of Digital Curation* 5 (1): 18–33. <https://doi.org/10.2218/ijdc.v5i1.141>.
- Blanke, Tobias, and Mark Hedges. 2013. "Scholarly Primitives: Building Institutional Infrastructure for Humanities e-Science." *Future Generation Computer Systems*, Special section: Recent advances in e-Science, 29 (2): 654–61. <https://doi.org/10.1016/j.future.2011.06.006>.
- Borgman, Christine L. 2010. "The Digital Future Is Now: A Call to Action for the Humanities." *Digital Humanities Quarterly* 3 (4). <http://www.digitalhumanities.org/dhq/vol/3/4/000077/000077.html>.
- Borgman, Christine L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, Massachusetts: The MIT Press.
- Boyle, Alan. 2018. "Crowdsourcing Efforts Save Digital Ghosts of Artifacts Lost in Brazil's Museum Fire." *GeekWire*. September 5, 2018. <https://www.geekwire.com/2018/crowdsourcing-efforts-save-digital-ghosts-artifacts-lost-brazils-museum-fire/>.
- Buchanan, Rose, Keau George, Taylor Gibson, Eric Hung, Daria Labinsky, Diana E Marsh, Richelieu, et al. 2021. "Toward Inclusive Reading Rooms: Recommendations for Decolonizing Practices and Welcoming Indigenous Researchers." *Archival Outlook*, Jan/Feb: 4-5, 14.
- Carpenter, Brian. 2019. "Archival Initiatives for the Indigenous Collections at the American Philosophical Society." *Case Studies on Access Policies for Native American Archival Materials*, no. 1: 1-13. https://www2.archivists.org/sites/all/files/Case_1_Archival_Initiatives_for_Indigenous_Collections.pdf.

- Christen, Kimberly. 2011. "Opening Archives: Respectful Repatriation." *American Archivist* 74 (1): 185-210.
- Christen, Kimberly, Alex Merrill, and Michael Wynne. 2017. "A Community of Relations: Mukurtu Hubs and Spokes." *D-Lib Magazine* 23 (5/6). <http://dlib.org/dlib/may17/christen/05christen.html>.
- "Citizen Archivist Dashboard." National Archives. Last modified November 7, 2018. <https://www.archives.gov/citizen-archivist>.
- "Citizen Scanning." National Archives. Last modified October 23, 2019. <https://www.archives.gov/innovation-hub/scanning>.
- Cooper, D., and O. Y. Rieger. 2018. "Scholars ARE Collectors: A Proposal for Re-Thinking Research Support." Issue brief. Ithaka S+R. <https://doi.org/10.18665/sr.310702>.
- Cox, Joe, Eun Young Oh, Brooke Simmons, Gary Graham, Anita Greenhill, Chris Lintott, Karen Masters, and Jamie Woodcock. 2018. "Doing Good Online: The Changing Relationships Between Motivations, Activity, and Retention Among Online Volunteers." *Nonprofit and Voluntary Sector Quarterly* 47 (5): 1031-56. <https://doi.org/10.1177/0899764018783066>.
- Cunningham, Adrian. 2011. "The Postcustodial Archive." In *The Future of Archives and Record Keeping: A Reader*, edited by Jennie Hill, 173-90. London, UK: Facet Publishing.
- Dallwitz, Dora, John Dallwitz, and Susan Lowish. 2019. "Ara Irititja and Ara Winki in the Apy Lands: Connecting Archives to Communities through Mobile Apps on Portable Devices." *Archives and Manuscripts* 47 (1): 35-52. <https://doi.org/10.1080/01576895.2018.1547652>.
- Daines, J. Gordon, and Cory L. Nimer. 2011. "Re-Imagining Archival Display: Creating User-Friendly Finding Aids." *Journal of Archival Organization* 9 (1): 4-31. <https://doi.org/10.1080/15332748.2011.574019>.
- Daniels, Morgan G., and Elizabeth Yakel. 2010. "Seek and You May Find: Successful Search in Online Finding Aid Systems." *The American Archivist* 73 (2): 535-568. <https://doi.org/10.17723/aarc.73.2.p578900680650357>.
- Ellis, D. 1993. "Modeling the Information-Seeking Patterns of Academic Researchers: A Grounded Theory Approach." *The Library Quarterly* 63 (4): 469-86.
- Evans, Max J. 2007. "Archives of the People, by the People, for the People." *The American Archivist* 70, no. 2 (Fall/Winter): 387-400. <http://www.jstor.org/stable/40294576>.
- First Archivists' Circle. 2007. "Protocols for Native American Archival Materials," accessed June 23, 2021: <https://www2.nau.edu/libnap-p/protocols.html>.
- Flinn, Andrew, and Elizabeth Shepherd. 2011. "Questions of Trust (and Distrust)." *Archival Science* 11 (3-4): 169-74. <https://doi.org/10.1007/s10502-011-9169-7>.
- Genovese, Taylor R. 2016. "Decolonizing Archival Methodology: Combating Hegemony and Moving Towards a Collaborative Archival Environment." *AlterNative: An International Journal of Indigenous Peoples* 12 (1): 32-42. <https://doi.org/10.20507/AlterNative.2016.12.1.3>.

- Hedges, Mark, and Stuart E. Dunn. 2018. *Academic Crowdsourcing in the Humanities: Crowds, Communities and Co-Production*. Chandos Information Professional Series. Cambridge, MA: Chandos Publishing, an imprint of Elsevier.
- Holley, Rose. 2010. "Crowdsourcing: How and Why Should Libraries Do It?" *D-Lib Magazine* 16 (3/4). <https://doi.org/10.1045/march2010-holley>.
- Huvila, Isto. 2008. "Participatory Archive: Towards Decentralised Curation, Radical User Orientation, and Broader Contextualisation of Records Management." *Archival Science* 8 (1): 15-36. <https://doi.org/10.1007/s10502-008-9071-0>.
- Iacovino, Livia. 2015. "Shaping and Reshaping Cultural Identity and Memory: Maximising Human Rights through a Participatory Archive." *Archives and Manuscripts* 43 (1): 29-41.
- Jansson, Ina-Maria. 2017. "Organization of User-Generated Information in Image Collections and Impact of Rhetorical Mechanisms." *Knowledge Organization* 44 (7): 515-28. https://www.ergon-verlag.de/isko_ko/downloads/ko_44_2017_7_f.pdf.
- Jennett, Charlene, Laure Kloetzer, Daniel Schneider, Ioanna Iacovides, Anna Cox, Margaret Gold, Brian Fuchs, et al. 2016. "Motivations, Learning and Creativity in Online Citizen Science." *Journal of Science Communication* 15 (3): A05. <https://doi.org/10.22323/2.15030205>.
- Macri, Martha J, Victor Golla, and Lisa L Woodward. 2009. "J.P. Harrington Project: Academic and Community Participation." *Survey Reports, Survey of California and Other Indian Languages*: 1-13. <https://escholarship.org/uc/item/6h5411pt>.
- Marsh, Diana E. 2018. "Toward Inclusive Museum Archives: User Research at the Smithsonian's National Anthropological Archives." In *Defining the Museum of the 21st Century: Evolving Multiculturalism in Museums in the United States*, edited by Yun Shun Susie Chung, Anna Leshchenko and Bruno Brulon Soares, 129-42. Paris: ICOM/ICOFOM.
- Matienzo, Mark A., and Amy Rudersdorf. 2014. "The Digital Public Library of America Ingestion Ecosystem: Lessons Learned After One Year of Large-Scale Collaborative Metadata Aggregation." *International Conference on Dublin Core and Metadata Applications*, October, 12-23. <https://dcpapers.dublincore.org/pubs/article/view/3700>.
- Mugglestone, Lynda. 2005. *Lost for Words: The Hidden History of the Oxford English Dictionary*. 1sted. New Haven: Yale University Press.
- Nimer, Cory, and J. Gordon Daines. 2008. "What Do You Mean It Doesn't Make Sense? Redesigning Finding Aids from the User's Perspective." *Journal of Archival Organization* 6 (4): 216-232. <http://doi.org/10.1080/15332740802533214>.
- Lanxon, Nate. 2011. "How the Oxford English Dictionary started out like Wikipedia." *WIRED*. January 13. <https://www.wired.co.uk/article/the-oxford-english-wiktionary>.
- Liew, Chern Li. 2016. "Social Metadata and Public-Contributed Contents in Memory Institutions: 'Crowd Voice' Versus 'Authenticated Heritage'?" *Preservation, Digital Technology & Culture* 45 (3): 122-33. <https://doi.org/10.1515/pdtc-2016-0017>.

- Oldman, Dominic, and Diana Tanase. 2018. "Reshaping the Knowledge Graph by Connecting Researchers, Data and Practices in ResearchSpace." In *The Semantic Web – ISWC 2018*, 325–40. *Lecture Notes in Computer Science*. Springer International Publishing.
- O’Neal, Jennifer R. 2019. "From Time Immemorial: Centering Indigenous Traditional Knowledge and Ways of Knowing in the Archival Paradigm." In *Afterlives of Indigenous Archives: Essays in Honor of the Occom Circle*, edited by Ivy Schweitzer and Gordon Jr. Henry, 45–60. Hanover, NH: Dartmouth College Press.
- Orli, Sylvia, and Jessica Bird. 2016. "Establishing Workflows and Opening Access to Data within Natural History Collections." *Collections: A Journal for Museum and Archives Professionals* 12(2): 147–62. <https://doi.org/10.1177/155019061601200208>.
- Osborn, Carlyn, Victoria Van Hying, and Mark Hartsall. 2020. "By the People Announces New Staff Campaign." *Library of Congress Gazette*, May 29, 2020.
- Osborn, Carlyn. 2020. "Celebrating Another Year with By the People | The Signal." Webpage. September 14, 2020. <https://blogs.loc.gov/thesignal/2020/09/celebrating-another-year-with-by-the-people/>.
- Owens, Trevor. 2014. "Making Crowdsourcing Compatible with the Missions and Values of Cultural Heritage Organizations." In *Crowdsourcing Our Cultural Heritage*, edited by Mia Ridge, 269–80. Farnham: Ashgate.
- Palmer, Carole. 2004. "Thematic Research Collections." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell Publishing. <http://www.digitalhumanities.org/companion/>.
- Palmer, Carole L., Oksana L. Zavalina, and Katrina Fenlon. 2010. "Beyond Size and Search: Building Contextual Mass in Digital Aggregations for Scholarly Use." *Proceedings of the American Society for Information Science and Technology* 47 (1): 1–10. <https://doi.org/10.1002/meet.14504701213>.
- Palmer, Carole L., Teffeu, Lauren C., and Pirmann, C. M. 2009. "Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development." Dublin, OH: OCLC Research and Programs. <http://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf>.
- Poole, Alex H. 2013. "Now Is the Future Now? The Urgency of Digital Curation in the Digital Humanities." *Digital Humanities Quarterly* 7 (2). <http://www.digitalhumanities.org/dhq/vol/7/2/000163/000163.html>
- Powell, Timothy B. 2016. "Digital Knowledge Sharing: Forging Partnerships between Scholars, Archives, and Indigenous Communities." *Museum Anthropology Review* 10 (2): 66–90.
- Ridge, Mia. 2016. *Crowdsourcing Our Cultural Heritage*. Digital Research in the Arts and Humanities. London: Routledge. <https://doi.org/10.4324/9781315575162>.
- Rolan, Gregory. 2017. "Agency in the Archive: A Model for Participatory Recordkeeping." *Archival Science* 17 (3): 195–225.

- Roued-Cunliffe, Henriette, and Andrea Copeland Japzon, eds. 2017. *Participatory Heritage*. London: Facet Publishing.
- Scheir, Wendy. 2006. "First Entry: Report on a Qualitative Exploratory Study of Novice User Experience with Online Finding Aids." *Journal of Archival Organization* 3 (4): 49-85. https://doi.org/10.1300/J201v03n04_04.
- Shilton, Katie, and Ramesh Srinivasan. 2007. "Participatory Appraisal and Arrangement for Multicultural Archival Collections." *Archivaria* 63: 87-101.
- "Shared Stewardship of Collections." Smithsonian Institution, 2019, accessed June 23, 2021, <https://folklife-media.si.edu/docs/folklife/Shared-Stewardship.pdf>.
- Society of American Archivists. 2021. "African Studies Library at University of Cape Town Needs Your Help!," April 20, <https://www2.archivists.org/news/2021/african-studies-library-at-university-of-cape-town-needs-your-help>.
- Somerville, Mary M., and Dana EchoHawk. 2011. "Recuerdos Hablados/Memories Spoken: Toward the Co-Creation of Digital Knowledge with Community Significance." *Library Trends* 59 (4): 650-662. <https://doi.org/10.1353/lib.2011.0018>.
- Spanner, Don. 2001. "Border Crossings: Understanding the Cultural and Informational Dilemmas of Interdisciplinary Scholars." *The Journal of Academic Librarianship* 27 (5): 352-60. [https://doi.org/10.1016/S0099-1333\(01\)00220-8](https://doi.org/10.1016/S0099-1333(01)00220-8).
- Stevens, Mary, Andrew Flinn, and Elizabeth Shepherd. 2010. "New Frameworks for Community Engagement in the Archive Sector: From Handing Over to Handing On." *International Journal of Heritage Studies* 16 (1-2): 59-76.
- Thompson, Samantha. 2017. "Why Don't Archivists Digitize Everything?", accessed June 23, 2021, <https://peelarchivesblog.com/2017/05/31/why-dont-archivists-digitize-everything/>.
- Trace, Ciaran B., and Unmil P. Karadkar. 2017. "Information Management in the Humanities: Scholarly Processes, Tools, and the Construction of Personal Collections." *Journal of the Association for Information Science and Technology* 68 (2): 491-507. <https://doi.org/10.1002/asi.23678>.
- Unsworth, J. 2000. "Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?" *Symposium on "Humanities Computing: formal methods, experimental practice"* King's College, London. <https://johnunsworth.name/Kings.5-00/primitives.html>.
- van Hooland, Seth, Eva Méndez Rodríguez, and Isabelle Boydens. 2011. "Between Commodification and Engagement: On the Double-Edged Impact of User-Generated Metadata within the Cultural Heritage Sector." *Library Trends* 59 (4): 707-720. <https://doi.org/10.1353/lib.2011.0011>.
- Van Hyning, Victoria. 2020. "Finding By the People Transcriptions in the Library's Digital Collections | The Signal." Webpage. July 9, 2020. <https://blogs.loc.gov/thesignal/2020/07/finding-by-the-people-transcriptions-in-the-librarys-digital-collections/>.

- Vilar, Polona. 2015. "Information Behaviour of Scholars." *Libellarium: Journal for the Research of Writing, Books, and Cultural Heritage Institutions* 7 (1): 17–39.
<https://doi.org/10.15291/libellarium.v7i1.194>.
- Washington Post*. 2020. "Pandemic Escape: Volunteers Transcribe Sally Ride's Papers, Rosa Parks's Recipes, Walt Whitman's Poems." May 18, 2020.
<https://www.washingtonpost.com/history/2020/05/18/historical-documents-transcribe-volunteers/>.
- Wiedeman, Gregory. 2019. "The Historical Hazards of Finding Aids." *The American Archivist* 82 (2):381–420. <https://doi.org/10.17723/aarc-82-02-20>.
- Wolfe, Heather and Victoria Van Hyning. Forthcoming. "More content, less context: Rethinking Access." *Archives: Power, Truth, and Fiction*, edited by Alison Wiggins and Andrew Prescott (Oxford: Oxford University Press).

A Call to Action: User Experience & Inclusive Description

Faith Charlton, Christa Cleeton, Alison Clemens, Betts Coup, Zoë Hill, and Jessica Tai

Abstract: Archivists have long sought to center users in our work, including our descriptive practices and choices, but there has never been a professional community of practice that supports or advocates for the user experience work necessary to truly understand our users' needs. This has become all the more critical as the archival field has turned its collective attention to reparative and inclusive descriptive projects. However, this work is often done without fully understanding the impact legacy or new description might have on users, particularly those from underrepresented communities. This white paper situates this issue in its current professional and institutional contexts and advocates for a community of practice with specific recommendations and the necessary equitable and long-term institutional support for user experience studies centered on the creation and evaluation of archival description.

Introduction to the issue

Our focus is the longstanding issue within the archives profession of the need to center and engage users in archival description.¹ Specifically, this work is necessary in order to better understand the impacts and ensure the appropriateness of the inclusive and reparative description efforts in which archivists are increasingly engaging for the benefit of users, particularly those from historically marginalized and underrepresented communities. It is critical for archivists to build pathways for user engagement in inclusive description decisions and to embed user experience processes as part of routine practices and workflows.

At present, there is a lack of data on the real impact of inclusive description efforts—from feedback forms to harmful language statements, edited terms and processing and contextualizing notes—possibly hindering the potential and prioritization of such initiatives or demonstrating that some of these efforts are unhelpful or in fact creating or perpetuating harm. The establishment of guidelines and best practices for user and usability testing in this area, as well as the infrastructure and resources required to do this work, will result in user-centered descriptive practices that are rooted in data rather than assumptions. It will also result in description that is responsive to and reflective of a wide array of communities who create, are represented in, and use historical records.

¹ Throughout this paper, the authors use “we,” “our,” and “us” to express the opinions of this cohort of practicing archivists and center the recommendations in our findings and experiences.

Background & context

User & usability testing

Archivists have long sought to center users in our work, including our descriptive practices and choices, and have bemoaned the fact that as a profession we consistently fall short of this goal. The recently revised DACS Principles codified centering users as not just a desire but an expectation for archival work. Declaring that users are the fundamental reason for archival description, Principle 2 notes that in order “[t]o make wise choices about descriptive practices, archivists must develop and maintain an awareness of user needs and behaviors” (Describing Archives: A Content Standard 2020).

Usability and user studies have been carried out over the years relating particularly to encoded description and finding aid websites (see references), but there is a significant lack of regular, routine user experience testing or even standards or guidelines for how to implement this work into practice within descriptive contexts. Thus, as Betts Coup notes in her article, “[a]lthough archivists often seek to center users in their processes of description, these efforts are often based on guesses concerning what users will find useful. Very little research has been done to understand how users read and navigate archival description; there is a distinct disconnect between the intention of centering users and carrying out usability or user studies to understand user needs.” She notes that this reality is a result of multiple factors, including “a lack of resources and support for routinizing user experience work around finding aids, as well as a significant amount of education regarding the design and management of user experience studies” (Coup 2021).

Thus, archivists continue to create description in a silo, with many attempts to center users based only on our conceptions of researchers’ needs. In addition, the users we do occasionally center are limited to select groups, typically groups that exercise power based on systemic privilege within our contexts; at our institutions (Harvard, Princeton, Yale), this primarily includes white academics.

Inclusive & reparative description

Building on an existing discourse of critical analyses of description work, the past several years have seen a significant increase in an interest in and focus on inclusive description within the library and archives professions (Tai 2020, Sutherland 2020, Tang et al. 2018, Arroyo-Ramírez et al. 2017, Bolding et al. 2020). This includes current conscientious descriptive work as well as retrospective or reparative projects to revise legacy description as some of the many ways in which librarians and archivists can work to dismantle white supremacy and cultivate a social justice-oriented approach to their work (Caswell 2017). The topic of inclusive description is increasingly prevalent at conferences and in lecture series and webinars, and more and more individuals and institutions are engaging in this work, including developing feedback forms, creating harmful language statements, editing problematic terms, and adding contextualizing notes to finding aids.

Based on a 2020-2021 survey that the Archives for Black Lives in Philadelphia’s Anti-Racist Description Working Group sent out to the archival community about the *Anti-Racist Description Resources* they published in 2019 (Antracoli et al. 2020), a majority of the 50 survey respondents noted that they or their institutions were beginning to engage in this work and that they had incorporated the Resources’

guidelines in some way.² An informal poll that was taken during the “Implementing Programmatic Anti-Racist (Re)Description at Predominantly White Institutions” session at the Rare Books and Manuscripts Section (RBMS) 2021 Conference demonstrated that of the 60 attendees who participated, 55% were engaging in inclusive description work and 30% were planning to do so. Additionally, various individuals and groups have developed recommendations, guidelines, and thesauri for archivists to implement community-generated and anti-oppressive description.³

In terms of centering users, two key components emphasized by proponents of inclusive description work are community collaboration and a more expansive view of audiences or users of libraries and archives. Proponents of inclusive description note the importance of community self-description, the need to equitably work with communities who created or who are documented in historical records, and the importance of privileging the language communities and individuals use to describe themselves. This also means collaborating with a wide variety of users to ensure that inclusive description work is in fact reflective of lived experiences, helpful, and mitigating harm.

Centering users in this way is a practice grounded in representational belonging. As defined by Michelle Caswell, Marika Cifor, and Mario H. Ramirez, representational belonging “empower[s] people who have been marginalized by mainstream media outlets and memory institutions to have the autonomy and authority to establish, enact, and reflect on their presence in ways that are complex, meaningful, substantive, and positive to them in a variety of symbolic contexts” (Caswell et al. 2016). Typically, and as explained by Caswell, Cifor, and Ramirez, representational belonging is enacted by community archives, but this could potentially be accomplished by other institutions if these marginalized communities and peoples are given the chance to voice their needs for description and access.

Institutional & professional milieu

Increased attention to inclusive description work has dovetailed with institutions' and professional organizations' recent focus on DEI-related initiatives and discourse with the elevation of the Black Lives Matter (BLM) movement and uprisings after the murder of George Floyd. At our respective libraries, this has included the publication of BLM statements and statements in support of the Asian American and Pacific Islander community, the formation of DEI committees and working groups, the creation of DEI or anti-racist LibGuides, the creation of DEI and anti-racism staff positions, and the revision of strategic planning initiatives, including those related to collecting and digitization efforts, to name some examples.

However, there is often a disconnect between institutions' stated support for DEI or social justice-related initiatives and what is actually implemented, prioritized, and routinized in practice and the inroads that still need to be made, particularly the fact that library and archives professionals continue to be roughly 80% white.

² Survey report forthcoming.

³ Some guidelines, as noted by Tai (2020), include: Archives for Black Lives in Philadelphia: Anti-Racist Description Resources, the Chicano Thesaurus, the Densho Terminology Guide, the Gender, Sex, and Sexual Orientation Ontology, the GLAAD Media Reference Guide, Homosaurus, the Indigenous Peoples Terminology Guidelines for Usage, National Center on Disability and Journalism Disability Language Style Guide, the Power of Words Handbook, and the Subject Headings for African American Materials.

In addition to institutional efforts, DEI and social justice-related activities and efforts have also become more prevalent within professional circles over the past few years, including initiatives of existing organizations like the Society of American Archivists (SAA), particularly its Diversity Committee's report and summary of feedback from the Black Lives and Archives statement discussion (Wooten and Mora 2021), as well as the formation of new organizations, such as We Here, the Abolitionist Library Association, Archives for Black Lives in Philadelphia, and the Blackivists, to name a few. These organizations and efforts have increased archivists' awareness and involvement in DEI and social-justice related efforts, which they are increasingly incorporating into their work.

Recommendations

As demonstrated in the first two sections of this paper, the need to center users in descriptive practices is a complex, pervasive issue that extends beyond inclusive or reparative description work. Centering user experience is an imperative in our field, and the need to implement this as a professional priority is underscored by considerations pertaining to inclusive and reparative work generally speaking. Taking multiple aspects of the issue into account, we offer several recommendations for how archivists and their institutions can address the need to center users in archival description.

Building a community of practice

Usability work in archives does not yet benefit from a robust professional community of practice. We advocate for building such a community in this area, as doing so will assure that usability work can be carried out in a way that is supported by peer-to-peer and professional organizational conversations and resources and informed by standards and guidelines.

We advocate for the creation of interest groups within the profession focused on usability and user testing that focuses on descriptive content as well as design and functionality. We specifically believe that the archival community would benefit from the establishment of a User Experience section or task force with the involvement of multiple sections within the Society of American Archivists (SAA), and we hope to be part of those efforts. We feel it is critical that a community of practice focused on usability and user experience of archival description be recognized formally. Such a group, whether an official SAA section or a group created out of a partnership amongst multiple existing sections, would be able to advocate for this work and also provide resources that would assist with it (e.g., testing scripts, usability toolkits, professional communication channels). SAA sections that might be involved with such a group include Accessibility and Disability, Description, Issues and Advocacy, Encoded Archival Standards, and the identity-based sections.

This community, no matter its governing structure, should develop a set of principles to guide user experience testing in archives, particularly testing related to descriptive content as well as design and functionality. These principles should be accompanied by a toolkit of scalable, customizable user and usability testing approaches and script templates. These guidelines and toolkit would provide an entry point for those implementing user experience testing as part of their practice. It could be maintained and supplemented as testing approaches develop and change over time. The toolkit should include feedback from those who have used it, in order to inform future testing decisions by archivists at repositories of various sizes, budgets, and audiences. To best facilitate and manage the principles and

toolkit, as well as ensure a long-term focus on user experience work in archival practice, the authors feel an SAA section might be the best way to formalize and structure this community of practice.

The archival community places a strong emphasis on shared standards, guidelines, and best practices. Creating standards and guidelines to inform user studies within archives would give validity to user experience work within our field and would allow library workers to ground their user experience work in a larger interprofessional context, no matter the size or budget of their institution.

Individual implementation

Having standards of practice surrounding user testing will allow archivists to conduct this work at their institutions and serve as an advocacy tool for archivists pushing to prioritize this work at an institutional level. Archivists who are interested in or currently engaging in inclusive or reparative description work should seek to learn and incorporate strategies and best practices from the user testing field while also collaborating with user experience (UX) colleagues or those with UX expertise to advocate for user testing and engagement models that are respectful, equitable, and relational, rather than transactional.

Importantly, as archivists we should approach our work with an acknowledgement of our positions and experiences and normalize that we do not and will not know everything. We must instead allow other forms of expertise to inform our decision making, particularly through community consultation and user testing. We must approach our work from a position of cultural humility. As Jessica Tai explains, "In emphasizing co-learning through community engagement, collaboration and partnerships, cultural humility refocuses archivists to be fundamentally user-centered. A pivotal step in doing so is to normalize not knowing. Cultural humility prompts archivists to acknowledge that they will never have all the answers, therefore opening space for other voices, and allowing nontraditional forms of expertise to inform decision making" (Tai 2020). This mindset allows archivists to approach descriptive decision-making from a point not of knowing but of open-minded inquisitiveness regarding other perspectives and experiences, especially those of users and communities.

Institutional responsibility

The ability to routinely center users as part of description work cannot happen without institutional support. Institutions should prioritize user testing, and, similar to the individual level, approach this work from a position of cultural humility, with "self-reflection and self-critique." As Tai states, "cultural humility moves beyond the role of the individual, shifting responsibility to institutions to enact policy change, create public-facing documentation, and cultivate intentional shifts in organizational culture..." (Tai 2020). Conducting inclusive, wide-ranging user and usability testing as part of regular archival practice, particularly testing related to reparative and inclusive descriptive projects, is another way an institution can adopt a position of cultural humility.

Academic libraries tend to place a heavy emphasis on the importance and centrality of their users, particularly students. Libraries with nascent or developing user experience departments have become more prevalent over the last ten years; however, the existence of this type of infrastructure with the resources and staff needed to conduct this work is not universal and often not prioritized, even by well-resourced institutions.

For user testing to be conducted in a meaningful, sustainable way, there must be administrative buy-in and support. It is often the case that staff carrying out user experience testing at any given institution lack the necessary training and expertise to do so. Well-resourced institutions should ideally hire those with expertise in this area to dedicated positions or support staff training for those who lack such expertise. These institutions must provide staff with sufficient time to become fully trained and then offer both time and compensation to carry out testing. Users and consultants who are asked to participate in user experience testing should be recognized for their contributions, whether through financial compensation or explicit acknowledgement for providing their expertise and perspective. This recognition should be seen as a crucial part of personal and institutional relationship-building.

It is well known that institutions need to be significantly better in terms of hiring and retaining a diverse staff; however, it is important that institutions not overburden their BIPOC employees with DEI-specific or reparative work. Concerning descriptive work specifically, a balance must be struck between ensuring that there is a diverse set of perspectives from which description is created and not automatically placing the responsibility for often difficult redescriptive or descriptive projects on BIPOC staff. Inclusive and reparative projects, which can include emotionally difficult labor, particularly by those who have experienced or been impacted by racism, bias, and prejudice, must be shared across staff. Institutions also have an ongoing practice of hiring contracted, term archivists and other staff, including students, to carry out descriptive projects. This illustrates the current, short-term interest in descriptive work and lack of commitment to considering this work as an on-going, permanently supported, critical effort for libraries and archives.

It is also necessary to welcome and ensure the inclusion of the perspectives of a wide variety of users and communities engaged with and related to the materials we steward. This work must be grounded in the concept of representational belonging, with a hope to draw ties between inclusive description and the affective impact of community inclusion and representation. We hope that user experience work focused on archival description can spark meaningful, relationship-based collaborations and experiences among users, library staff, and collections, and we hope that user experience work will serve as a catalyst and vehicle to share and distribute institutional power.

Scalability

The authors recognize the importance of creating an approach to user testing that is scalable to ensure that this work can be carried out at various institutions ranging in size and budget and within different communities. While standards and guidelines help to inform this work, each institution will approach user testing in a way that is unique to their own needs and limitations. Because of this, user testing should be flexible enough to allow for multiple approaches, including informal and formal testing methods.

Still, having a welcoming, recognized community of practice that is easily accessible to archivists working at institutions of all sizes and budgets is necessary to ensure scalability. Archivists who might be lone arrangers or working in community archives lacking financial support could participate in conversations with others interested in this work, have the opportunity to use and customize testing scripts developed at larger, more well-funded institution, and participate in an ongoing discussion of how best to ensure user and usability work is inclusive and relationally-focused.

Conclusion and next steps

In order for libraries and archives to be as user-centered as they profess to be, user experience work must be included as part of regular practice, including in the area of reparative and inclusive description. The next steps for this work, as outlined above, are significant and sizable, requiring participation at individual, institutional, and professional levels, far beyond this small cohort. We need the voices of archivists who work at repositories of varying sizes and types to advocate for this work; we need the perspective of community members, groups, and staff members whose voices have been ignored in our largely white-centered (descriptive) practices; and we need institutional and professional support to develop standards and implement user experience studies as part of everyday archival work. To do this in a collaborative, scalable, and relational way with a wide range of users and communities, particularly those from historically marginalized groups, with whom we have not traditionally collaborated, it is critical that we create a formalized, accessible community of practice in order to share ideas, establish standards and methods, advocate for resources, and continue the conversation we hope this paper will spark.

Acknowledgements

We would like to thank the organizers and facilitators of the Lighting the Way Working Meeting, for providing the space and structure for our conversations that led to this white paper. In particular, we would like to thank our group's facilitator, Max Eckard, for his assistance and feedback.

References

- Antracoli, Alexis A., Annalise Berdini, Kelly Bolding, Faith Charlton, Amanda Ferrara, Valencia Johnson, and Katy Rawdon. 2020. "Archives for Black Lives in Philadelphia: Anti-Racist Description Resources." October.
https://archivesforblacklives.files.wordpress.com/2020/11/ardr_202010.pdf
- Arroyo-Ramírez, Elvia, Molly Brown, Dinah Handel, Jasmine Jones, Shannon O'Neill, Rachel Mattson, Giordana Mecagni, Holly Smith, and Kelly Wooten. 2017. "Radical Empathy in Archival Practice." Presentation, Society of American Archivists Annual Meeting, Portland, OR, July.
<https://archives2017.sched.com/event/ABGy/301-radical-empathy-in-archival-practice>
- Bolding, Kelly, Faith Charlton, Betts Coup, Cate Peebles, Jennifer Garcon. 2020. "Implementing Inclusive (Re)Description at Predominantly White Institutions." Presentation, Society of American Archivists Annual Meeting, August.
- Caswell, Michelle. 2017. "Teaching to Dismantle White Supremacy in Archives," *The Library Quarterly* 87, no. 3 (July): 222-235, <https://doi.org/10.1086/692299>.
- Caswell, Michelle, Marika Cifor, and Mario H. Ramirez. 2016. "To Suddenly Discover Yourself Existing: Uncovering the Impact of Community Archives," *The American Archivist* 79 (1): 56-81.
<https://doi.org/10.17723/0360-9081.79.1.56>
- Coup, Betts. 2021. "The Value of a Note: A Finding Aid Usability Study" *Journal of Contemporary Archival Studies* 8, Article 13. <https://elischolar.library.yale.edu/jcas/vol8/iss1/13/>.

- Daniels, Morgan and Elizabeth Yakel. 2010. "Seek and You May Find: Successful Search in Online Finding Aid Systems," *The American Archivist* 73 (2): 535-568, <https://doi.org/10.17723/aarc.73.2.p578900680650357>
- Society of American Archivists Technical Subcommittee on Describing Archives: A Content Standard (SAA TS-DACS). 2020. *Describing Archives: A Content Standard*. Version 2019.0.3. Chicago: Society of American Archivists. (https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf)
- Hintz, Carrie and Sarah Quigley. 2021. "A Matter of Trust: Practical Strategies for Writing User-Centered, Values-Driven Description," *Journal of Archival Organization*. Published ahead of print, June 13, 2021. <https://doi.org/10.1080/15332748.2020.1806615>
- Nimer, Cory and J. Gordon Daines. 2008. "What Do You Mean It Doesn't Make Sense? Redesigning Finding Aids from the User's Perspective," *Journal of Archival Organization* 6 (4): 216-232, <https://doi.org/10.1080/15332740802533214>
- Sutherland, Tonya. 2020. "Redescription as Restorative Justice" LDL Speaker Series (November) https://www.youtube.com/watch?v=xOs_r6D9lNk
- Tai, Jessica. 2020. "Cultural Humility as a Framework for Anti-Opressive Archival Description," *Journal of Critical Library and Information Studies* 3 (Vol. 3: Radical Empathy in Archives, 2020) <https://journals.litwinbooks.com/index.php/jclis/article/view/120>
- Tang, Annie, Dorothy Berry, Rachel Winston, and Kelly Bolding. 2018. "Toward Culturally Competent Archival (Re)Description of Marginalized Histories." Papers presented at the Society of American Archivists Annual Conference, Washington, DC, August. <https://archives2018.sched.com/event/ESld/101-toward-culturally-competent-archival-redescription-of-marginalized-histories>
- Walton, Rachel D. 2015. "Looking for Answers: A Usability Study of Online Finding Aid Navigation," Master's Paper, University of North Carolina. <https://doi.org/10.17615/6wkj-4654>
- Wooten, Kelly and Teresa Mora, with the Society of American Archivists Diversity Committee Members. 2021. "Diversity Committee: Feedback and Recommendations Following Black Lives and Archives Forum," Society of American Archivists Council Meeting, March 17. <https://www2.archivists.org/sites/all/files/0321-V-B-DivComm-BlackLivesForumFeedback.pdf>
- Yakel, Elizabeth. 2004. "Encoded Archival Description: Are Finding Aids Boundary Spanners or Barriers for Users?" *Journal of Archival Organization* 2 (1-2): 63-77. https://doi.org/10.1300/j201v02n01_06

Speeding Towards Remote Access: Developing Shared Recommendations for Virtual Reading Rooms

Elvia Arroyo-Ramírez, Annalise Berdini, Shelly Black, Greg Cram, Kathryn Gronsbell, Nick Krabbenhoeft, Kate Lynch, Genevieve Preston, and Heather Smedberg

Abstract: Virtual Reading Rooms (VRRs) have the potential to provide secure, remote online access to digital archival materials but VRRs have traditionally been developed as highly customized, siloed systems informed by hyper-local requirements and inconsistent or unclear commitments to sustainability. This report presents considerations and recommendations for the holistic development and sustainability of VRR systems as informed by common themes and challenges associated with existing VRR systems at each of the authors' institutions. It is the hope of the authors that these recommendations will help inform a set of principles for VRR stewardship that take into account user needs; allocation of resources; policy creation; copyright and ethical needs for future development of VRRs; and responsible stewardship of mediated digital discovery and access to archival collections.

Introduction

An increasing number of cultural heritage institutions are taking steps to provide secure, remote online access to their digital archival materials. Several institutions have developed their own "Virtual Reading Rooms" which consist of specialized, often customized systems that enable mediated online access to users.

The term "Virtual Reading Room" (referred hereafter as VRRs) can be used to describe a number of user-centered approaches to remote access (Research Libraries UK 2021). For the purposes of this document, the authors have chosen to focus on the Society of American Archivists' definition which defines a VRR as a controlled online research environment in which the discoverability and/or downloading of archival materials is deliberately limited, and/or access is restricted to users who have created a limited-term account and agreed to terms of use similar to those that apply in a physical reading room of an archival repository (Society of American Archivists 2021).

Interest in the area of remote access has been significantly heightened by the onset of the coronavirus pandemic, which has forced libraries, museums, archives, and other cultural heritage institutions to shutter their onsite operations, impacting their ability to service their materials in person. At the same time, access and discoverability of digital archival materials are enduring challenges even as various institutions return to onsite service models; it is imperative the cultural heritage profession continues

to seek thoughtful, sustainable solutions towards holistic development and practice of VRRs that account for people, policies, technologies, and workflows that cut across traditional department lines, and areas of responsibility.

Background

The authors include IT developers, copyright specialists, public services librarians, and archivists who have experience in the development and practice of VRRs at their respective institutions. The institutions they represent include U.S. based universities, government, public, and private sector institutions.

Findings from the *Lighting the Way: A Preliminary Report on the National Forum on Archival Discovery and Delivery* indicated a strong interest in the area of VRRs during the Lighting the Way Forum. Several authors of this paper participated in discussions during the Forum that involved possibilities for a national, consortial, or otherwise shared VRR service; developing a list of copyright variables relevant to fair use and VRRs; engaging with existing software services; and establishing communities of practice specifically for VRR development and sustainability (Matienzo et al. 2020).

For the Lighting the Way Working Meeting (2021) the authors proposed to explore answers to the following questions:

What are the system agnostic, minimum requirements to develop a VRR system and service? What human, intellectual, and material resources are needed for both system development and ongoing service? What are the legal and technical considerations that will underpin the development of VRRs?

Throughout our exploration of these questions the authors were guided by a shared working purpose statement: “Articulate commonalities and differences in development and use of VRRs, prioritize cultural sensitivity in those activities, with both of these actions in service of enabling sustainable and equitable access.” This purpose statement has driven the deliverables herein.

Purpose

The primary audience for this report is library and archives practitioners, administrators, and copyright experts who are interested in learning about system agnostic, minimum considerations to develop VRR systems and services. Target individuals include archivists, public services staff, catalogers, curators, software developers/maintainers, library and/or archives administrators, rights holders (donors, records creators, Indigenous communities), and lawyers. Researchers and general members of the public may also find this report useful.

The report presents common themes and challenges associated with developing and sustaining VRRs as informed by the authors’ local VRR development and practices. It presents considerations and recommendations for the holistic development and sustainability of VRR systems which, the authors hope, will help inform a set of principles for VRR stewardship that take into account user needs; allocation of resources; policy creation; and copyright and ethical needs for future development of VRRs. The authors’ goals for this report are 1) to serve as a starting point to break away from the siloed approaches to VRRs in both system development and practice; 2) to guide future exploration towards

communities of practice for VRRs; and 3) provide a foundational, but in no way definitive, list of best practices towards responsible stewardship of mediated digital discovery and access to archival collections.

VRRs in the Ecosystem of Access and Discoverability

VRRs have the opportunity to play a key role in the stewardship and access of archival collections with mediated access needs. Two primary barriers to providing greater online access to archival collections are the increasing volume of born-digital and digitized materials; and the fact that by their nature, many archival materials in physical and digital formats preclude archivists from sharing them freely online even if they had the logistical capacity to do so in a VRR. A mediated VRR access model can lower both barriers to access. Legal considerations for requiring mediated access include restrictions due to copyright, personally identifiable information (PII), protected health information (PHI), and donor-imposed restrictions. Other motivations for a mediated approach seek to address equitable access goals such as limiting broad online access to culturally sensitive archival materials; respecting community partnership agreements; and the ability to provide access to materials that cannot be made openly available online without requiring users to travel to the reading room.

Internally, VRRs have the potential to reduce impact on access services staff who may spend large amounts of time downloading, uploading, or sharing files from their DAMS systems to fulfill individual research requests through third-party cloud storage solutions or to a dedicated workstation in the reading room.

Though the potential for VRRs to effectively plug a gap in the access paradigm is becoming increasingly understood, these systems have, thus far, been developed as highly customized and siloed systems informed by ad-hoc, hyper-local requirements. Additionally, these systems have been developed by institutions who have devoted large amounts of resources to their initial development and share largely inconsistent approaches to sustained resources. Grounded in our collective experience representing institutions with varying deployments of VRR systems and practice, the authors would like to warn of technology solutionism, which is a tempting option in the current landscape of VRRs. Technology solutionism proposes that all issues—from the personal to the political—have technological ‘fixes.’ Implicitly, this assumes that all challenges (historically, presently, and in the future) can be resolved through the use of technology, often regardless of human, climate, or other types of cost (Selinger 2013).¹

Ultimately, we observe and recognize that a technology-as-savior approach perpetuates issues with vendor influence in the cultural heritage profession, reliance on other (often commercial) industries’ impact on the direction of technology development, and temporarily encourages unsustainable field and organizational practices by providing short-term “wins” through early success with technology-first solutions for operational work.

While this challenge is not unique to VRRs, there is a pattern across cultural heritage institutions that points to a precarious reliance on custom, specialized, or legacy tools to serve as ongoing access points

¹ “We continue to expect technology to deliver us from the imperfections of the human condition, though history doesn’t support that idea.” – Evan Selinger (2013), summarizing a key concept from Evgeny Morozov’s book *To Save Everything, Click Here: The Folly Of Technological Solutionism*.

without responsibly-developed sustainability plans for the tool or its dependencies (technical and administrative). To that end, this report attempts to demystify technology considerations through naming and discussion around the roles played by various aspects of technology and human interaction with technology.

To date, there has never been a list of requirements or considerations to help inform the development or sustainability of VRRs in the professional literature. For a profession struggling with precarious labor and funding, coupled with the evolution of user behavior and expectations, the cultural heritage profession stands to benefit from community knowledge building around planning, implementation, growth, and maintenance of VRRs.

The technological development recommendations and service considerations outlined in this report are one step toward building a community of practice for VRRs. They can be used to create a roadmap for improving access, increasing engagement with collections, addressing the needs of the user in a virtual setting—all while centering human-first priorities like cultural sensitivity and equitable access. Responsible implementation and sustainability of VRRs require many considerations such as resources of funding and staff; advocacy and outreach for the creation and sustenance of the VRR; and ethical considerations relating to culturally sensitive knowledge, legal restrictions, privacy of donors, and creators and users. The desired result is to ethically and equitably provide secure and mediated VRR access to collections not well-suited for open access models.

Advocacy and Outreach

VRRs need ongoing funding, staff, and time in order to be built and maintained sustainably. The first and sometimes most challenging step in VRR development is advocating for these resources to administrators, who have myriad priorities to balance, or may view a VRR as a one-time project. Requesting support and prioritization for a VRR can be situated within advocacy for expanding access to archival materials, having a comprehensive understanding of the legal risks (see **Copyright** section), and better support for stewardship of born-digital and digitized materials overall. Other routes for advocacy include highlighting the equitable access VRRs provide to users, opportunities for more culturally sensitive access approaches, and cutting-edge methods.

VRR development requires input from stakeholders, resources for support and development, and continual assessment and evaluation of whether goals are being met or improvements are needed. To achieve these goals, obtaining buy-in from stakeholders within and outside of the library/archive is key. Begin with outreach to the administration to advocate for the service and the resources needed to sustain it, and continue outreach to partners and users to ensure that the system is fulfilling its intended purpose. Advocacy and outreach for VRRs is ongoing and should be considered part of regular maintenance procedures. Opportunities to share knowledge across professional groups, conduct or contribute to research on current VRR models, how VRR systems can reach more people, and integrate with other systems, are examples of ways staff can help perpetuate VRRs as a core method of access.

Systemic issues in the profession regarding digital preservation and access have a direct impact on VRR development. Recent studies of digital preservation practitioners show “even as the field of digital stewardship is entering a period of operational maturity, practitioners largely consider digital stewardship values and goals to be misunderstood at an organizational level” (Blumenthal et al. 2020).

Perceptions about the needs for digital preservation and access between practitioners and administrators can be significant, as manual or ad-hoc processes managed and implicitly performed by practitioners are understood or accepted as common practice whereas these practices can be significantly improved by the existence of comprehensive support for digital preservation and access programs determined by administration and leadership.

VRRs are complex access systems that larger institutions often necessitate cross-department collaboration and buy-in. Smaller institutions may have one position or department carrying out many of these roles, so cross-department collaboration may not be possible. In these cases, those advocating for VRRs may need to focus more outreach to boards or funding groups that will be less intimately knowledgeable with processes, workflows, and benefits. Building a strong partnership based on collaboration and shared understanding is crucial. It is important to understand that developers likely have many other commitments, and that adding a VRR will expand their portfolio over time, and not only during the development phase. Archivists or others championing for developing a VRR will need to effectively advocate for why this work needs to be prioritized among the many other demands on their time. Including archival and technical services staff and public services staff during early stage outreach and advocacy efforts can help guide the development of the VRR functional requirements in useful ways and can help describe specific, tangible benefits to investing in a VRR. Some organizations may also benefit from outreach to a records management team, research data management unit, and general counsel. Counsel may be particularly important to consult, considering the potential copyright and legal implications of making content more widely available.

Engaging with the research community stakeholders on a regular basis is critical to the VRR. Users are a stakeholder group that should be consulted both during and after development to see what is desired, useful, or in need of improvement. Usability considerations should be baked into the process whenever possible. Education and communication should be the cornerstone of user outreach that will not end with development but be part of the launch of the VRR (see **Equitable Access**).

Along with general users, consider donors, Indigenous and/or other community groups among the stakeholders who will help inform your requirements, and may help you advocate for the need for this level of service (see the following section, **Ethical Concerns: Culturally sensitive materials**).

Recommendations: Advocacy and Outreach

- *Gain administrative support for developing VRRs by emphasizing their legal basis (see Recommendations: Copyright section), and popular benefits of VRRs like secure equitable expanded access, cutting edge technology, and more culturally sensitive access methods. Alternatively, identify institutional strategic priorities (e.g. improve discoverability and use of archival materials) that align with a VRR and build a case around those priorities*
- *Be prepared to articulate the importance of granular access options for digital archives, and to describe archives workflows and practices more generally, when speaking with external systems developers, higher level administrators, general counsel, users, donors, and other key stakeholders*
- *Engage the necessary stakeholders at the onset of developing the VRR and include software developers, access services staff, technical services, records management team, research data*

unit, general counsel, users, and communities with culturally sensitive needs throughout the development process

Resources

It is essential for an organization planning to develop a VRR to first recognize that it is a service, not a project. VRRs require sustained funding, people, time, and ongoing research and maintenance. These systems cannot be static as they must accommodate new file formats, changing restrictions, and legal concerns over time. In addition, VRRs are a new service that will require both experimentation and skill development amongst staff tasked to manage its technological and service needs. Current iterations are only the first in the long road of what VRRs may become. Support for staff to do this work, and the staff to do this work, is key.

Funding

While there are many resources that factor into the development of a VRR, at the foundation is funding. Funding for developers, devoted staff time, vendor contracts, maintenance costs, digitization costs, and a myriad of other needs all require institutions to commit adequate and appropriate budgets to make a VRR not only functional, but lasting. Across the represented VRRs referenced herein, each required dedicated funding for staffing and staff time, vendor contracts, other costs, or all of the above (see Appendix A table).

At its heart, funding for VRRs comes back to the overall costs of stewardship for collections. It is a new form of access that will become essential with the increase in born-digital and digitized materials in collections, and as users seek alternative methods of access to traditional reading rooms. OCLC's *Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections* report points out that born-digital collections are increasing in size and frequency, and "require specific equipment and expertise in order to preserve and provide access to them, or time to experiment and problem solve" (Weber et al. 2021, 7). "Specialized equipment," which can be interpreted to include specialized software and services like VRRs, are not one-time costs, and must be factored in as regularly occurring and essential pieces of any institution's budget.

Staffing

Staffing is one of the most important components requiring funding and is often neglected (Weber et al. 2021, 11), both in hiring for new expertise or building in capacity for existing staff. For example, Princeton University Library's VRR, an in-house solution built on an existing access system, required a team of seven staff members meeting bi-weekly over the course of a year (and ongoing). This did not include the time IT put into development work. Time and money were taken from other projects, but the work was prioritized due to the urgent need for remote access to content caused by the coronavirus pandemic. Princeton University Library's VRR is now a service that must be maintained and is resourced to be iteratively updated. Without dedicated staff continuing to work on it, further development and improvement would cease.

Ideally, VRR costs should be earmarked as part of a dedicated digital access/preservation funding pool, to ensure that the monetary support for these roles, hours, and other resources does not dry up or become devoted to another priority. Funding for the VRR and its staffing should not come from

temporary pools or endowments that cannot sustain a VRR in perpetuity. VRR systems and service models are still nascent, and staff should also engage with research opportunities to build communities of practice for VRR. Opportunities to share knowledge across professional groups, conduct or contribute to research on current VRR models, how VRR systems can reach more people, and integrate with other systems, are examples of ways staff can help perpetuate VRRs as a core method of access.

Software Developers

Currently, many common VRR technological needs require some level of software development expertise in-house. There are no out-of-the-box solutions or modules that build on commercial repository software packages that meet these needs, and the requirements of VRR needs are extremely nuanced per institution based on the very specific nature of donor agreements, handling conditions for materials, nature of materials, and technological limitations of individual reading rooms. Software developers will be responsible for the build and maintenance of the system, ongoing development of add-ons, plug-ins, and connections to other services, as well as training and education for staff of the initial workflows.

Software developers often have the best sense of how much time will be needed for the initial development and what future steps are necessary to consider to continue sustaining an in-house VRR system. The required effort varies project to project, and will be impacted by the platform(s) in use to serve digital content, and how easily these platforms can be extended programmatically.

Maintenance

VRR workflows require ongoing user account mediation and review. Users and user groups may be granted access on a term-limited basis, or need timely access, and then need to be removed in a timely fashion. Maintenance includes ensuring that programmatic access reflects legal recommendations which is vital to creating a service that is supported long-term. It can be beneficial for the workflow to be developed as a shared process between special collections and repository management staff to ensure that individual access is mediated appropriately, and that those who developed the service technically and those who maintain the service from a user perspective are on the same page in terms of the growth/scope of its use.

Public Services Staff

Reading room, reference, and instruction staff will be, oftentimes, the public face of the VRR and its use. They will regularly build public awareness and train users on this service and how it fits in with the broader service model for archival and digital materials. This group may also be responsible for granting user access and is also the staff at the front line of troubleshooting users' challenges with navigating the discovery, access, and use of the system. Public services staff will need to understand the system well enough to take on troubleshooting and report problems appropriately. As they have direct expertise with understanding users' needs, it would be wise to include these staff in early development conversations, and encourage ongoing feedback from this cohort toward enhancements. It is absolutely crucial to keep these staff well-informed about changes.

Archival Processing, Metadata, and Technical Services

Staff responsible for processing and creating metadata for digital collections also have important roles to play, as VRRs will serve as a new endpoint for the work they do to provide discovery and access to archival materials. These groups typically create necessary workflows, are responsible for the ingest of new materials and metadata to VRR systems, and draft collection guides like finding aids that help users discover materials on the VRR. Digital archivists often play key roles in the development of functional requirements, especially where VRRs are designed to create access to born-digital records. While the role of systems developers has already been well-described, staff who carry out metadata work within digital collections workflows are also crucial to both development and ongoing maintenance of the system, particularly when a VRR is built on an existing digital collections infrastructure.

Curatorial

Staff in curatorial roles will contribute to ongoing decisions for collections to serve via the VRR, as further digitization or acquisition takes place. They will periodically evaluate when collections may be liberated from mediated process (e.g. copyright or donor restrictions sunset), or as open access licensing agreements take effect. Like public services staff, curators may also serve as the public faces of the VRR service. They will likely incorporate information about VRR service in donor conversations about access. While curators may not need to know all the technicalities of the system, they will need to be well versed in its overall operation, and have a line of communication to the system developers and to seek enhancements down the road as their collection, donor, or user needs evolve.

Time and Professional Development

Born-digital materials, in general, require extra staff time for experimentation, writing documentation, testing, reworking, re-testing, and communication (Weber 2020). This often includes processing, but does not always accommodate designing an entirely new access system for those materials and for digitized content. In order to create a successful service, staff will need time to establish the functional requirements to build a VRR, format collections for ingest, add necessary description or rework finding aids to connect to the system, develop new features, research and consult with stakeholders regarding restrictions and copyright concerns, and may need to go back and change settings that did not work after testing. VRRs are still new and most current iterations are one-offs, custom systems for individual institutions. At best, staff working on designing a VRR that will work for them will need time to research options, ask other institutions how they established a similar program, and investigate institutional roadblocks and legal concerns.

Recommendations: Resources

- *Establish permanent funding source for VRRs before development of the system and service*
- *Ensure adequate permanent staff to maintain and develop VRR; developing and sustaining a VRR system will require some level of in-house software development expertise; consider adding language to job descriptions or annual plans specific to VRR system development and service management*

- *Build in specific time to execute the development of the VRR and account for the professional development needs (skill building, knowledge sharing, research) staff will need to initiate and maintain a successful service*
- *Ensure all staff included in discussions and fact gathering for development of system; including software developers, maintenance, public service, reference, instruction, digital archivists and curatorial staff*

Users and Use of VRRs

When developing technological requirements for VRRs, most institutions will have different categories of users they serve. Academic institutions may grant different access privileges to their campus affiliates vs. the general public, while public libraries/archives/museums may not have such distinctions. Public libraries may want to provide different access levels to staff vs. public users. Some digital collections are available freely on the web, while others benefit from mediation; an individual user may themselves, at times, fall into different access categories for different materials. The level of granularity of these use cases will vary from place to place, granular access management capability is a consistent need across VRRs.

The biggest common programmatic need for VRRs is the ability to reasonably, and judiciously, limit or contain access to restricted materials at a granular level. If possible, it is often considered good practice to programmatically block the download of items by the user, a good-faith effort can be sufficient. There is no way to prevent users from taking screenshots/pictures on their mobile devices of items in the VRR environment, but software can generally programmatically disable download and “right-click/save as” functionality for embedded media in the VRR player. Time-based access is an option for users to access the material for any length of time within a date range depending on VRR policies and need. Another option is user-limited access, a condition wherein a set number of simultaneous users may access the material to simulate reading room conditions; for example, a set of users who might normally view the material as part of an instructional session with special collections, or a couple of users working together simultaneously on a group project.

Common granular access use cases can be programmed into digital content-serving platforms present in broad categories of technical functionality. They can be used in combination to meet a variety of VRR needs and policies. These include authentication, where users must be a member of an organization, acquire a sponsored institutional guest account if they are an approved user but are not affiliated with the organization, or have a local account with a non-organizational authentication scheme created for them in order to log in. This last method should only be used if in compliance with the organization’s VRR access policy, and only if the first and second methods are not achievable. Access can also be limited based on location, through methods such as IP restriction where users from within a specified set of IP ranges can access material. This can include individuals on-campus as well as VPN users whose remote access over the VPN connection simulates being on-campus, and could include fixed IP addresses from remote campus locations. Click-through agreements, where no login is required but users must complete an action to assert that they have read and understand the terms of use before accessing material, is another common case for materials with conditions for use separate from or in addition to access. Additional object- or collection-level restrictions may apply, such as limiting the

amount of time during one session that a user has to access an object in the VRR environment, or applying embargo functionality, releasing objects for access after a specified date.

Privacy of VRR Users

Naming and taking actionable steps towards the safety, security, and autonomy of end users should be a critical focus when developing VRR systems and services. Acknowledgement and prioritization of end user privacy demonstrates an ethical responsibility to our current and future audiences. Examples of this can include transparent and understandable privacy policies for websites and online products, clearly worded terms of service or agreements, and any effort to demystify what an organization is doing internally with data collected about users.

Institutions need to address their role in this user data creation, consumption, and distribution ecosystem. At an intersection with library patron protections, policies and practices around tracking, analytics, and/or monetizing or leveraging user data are often obscured, intentionally or benignly, through things like legal jargon or extremely long agreements with click-through acceptance actions. VRRs are especially subject to these risks because, often, users are asked to perform actions that expose personal data like:

- verify their identity,
- provide detailed information about research needs and credentials, and/or
- are monitored in systems (views, clicks, downloads, exports, etc.) including integrated or third-party software, which may have its own complex tracking and reporting functions for user behavior.

User data rights can be ignored in an effort to gain desirable usage and engagement metrics for the institution. This incentivizes ignoring responsible end user care in order to demonstrate the value of the system—a dangerous precedent that is further pushed by limited awareness of current global and domestic user rights outside of legal or front-end web teams responsible for compliance work.

Accessibility

Another area of acknowledging users as core agents of our service is accessibility of virtual environments. As of 2021, the Americans with Disabilities Act (ADA) does not define "web accessibility" requirements. Instead, the W3C's Web Content Accessibility Guidelines (WCAG) are used for evaluation and compliance in certain businesses. The guidelines are "stable, referenceable technical standards [...] with a goal of providing a single shared standard for web content accessibility that meets the needs of individuals, organizations, and governments internationally" (W3C 2021). The WCAG identify 3 levels of conformance:

- A (bare minimum level of accessibility)
- AA (target level of accessibility meeting legal requirements)
- AAA (exceeds accessibility requirements)

These conformance levels—used by institutions and businesses to audit their custom and commercial technologies, policies, and user experiences—address the four Principles of WCAG: Perceivable; Operable; Understandable; Robust. These guidelines are codified into an ISO standard (ISO/IEC 40500:2012), reflecting a global need for guidance and structure around ensuring accessibility in virtual environments.

Accessibility deficits are extremely evident in virtual offerings from U.S.-based archives, museums, and libraries. There is a noticeable increase in professional attention and funding related to remediating accessibility failures. Accessibility challenges will continue to multiply, layer, and perpetuate inequality in the way collections are made available generally. Centering accessibility is a way forward for all audiences—it reflects an acknowledgment of user expectations and abilities, prioritizing the human participants in our attempts to disseminate material and knowledge.

Equitable Access

VRRs have the potential to allow repositories to reach wider audiences beyond a physical reading room. They provide remote access for those who cannot or do not want to travel around the world to do research, and they are an alternative for those who do not wish to enter spaces that may be hostile to their presence (Farmer 2018). It is important to note here that providing new methods of access to our archival materials through VRRs does not solve the problem nor absolve the cultural heritage profession from naming and eradicating exclusionary reading room practices, racist collection development practices, prejudice, and mistreatment that persists in archives. If not well informed by the needs of donors, creators, users, subjects, and communities, VRRs may perpetuate the same exclusionary practices that have been reported in our physical reading rooms.

Implicit issues with VRR access models include the acknowledgement that not all users will have access to fast internet speeds, comparable internet speeds, or any internet at all. This could lead to a difficult or impossible virtual research experience. Loading large, high resolution images or audiovisual files could be problematic. Splitting up large file sets may be one way to address this problem, as well as providing metadata about file download sizes and download time estimates to open materials so that they can choose what to download and when. Where possible and appropriate, providing VRR access to low resolution access copies or more common file formats may be preferable to original preservation formats.

Engaging with the research community on a regular basis, providing announcements on the VRR and how to use it, and performing frequent evaluations to identify access gaps has to be part of the lifecycle of any VRR. Education and communication should be a cornerstone of the launch; not every user will feel comfortable navigating through the steps required for access to mediated content, and not every user has the technological capabilities to learn how to use the system on their own. Having staff be available to answer questions, provide training, and walk through use of the VRR as researchers learn of its availability will not only improve the user experience, but it will also let more people know that the service exists.

Recommendations: Users and Use of VRRs

- *Establish user access requirements, listing needs for various groups of users*
- *Establish protocols for restricting access, preventing downloading and time-based access*
- *Program into platform access based on VRR policies; including authentication of users, limitation based on location, and click through agreements*
- *Establish policies and protocols to secure user data collected for using the VRR*
- *Frequently engage with the research community to learn how they use the VRR and how well it meets their needs. Evaluate gaps and modify the service to reach the most users*
- *Include an easy to use feedback mechanism so that users can independently reach out with comments. Create frequent training and marketing opportunities so that users understand the VRR and how and when to use it*
- *VRRs must work with assistive technology (UC Berkeley 2021). Screen readers, text to speech apps, screen magnification software, and other assistive technologies should not be an afterthought in development. Many institutions have minimum web access accessibility requirements (Princeton University 2021) that can help guide project teams through development*
- *User testing should be done on a regular basis to make sure that the technology actually works and continues to do so*
- *There may already be an accessibility team that can help evaluate a new web service at your institution, system developers may be familiar with best practices, and users can tell your institution when something does not work for them*

Ethical Concerns

The implementation of VRRs requires ethical considerations relating to: culturally sensitive knowledge; privacy and confidentiality of donors, creators, and users; and accessibility. These areas may appear daunting, as they require nuanced decisions balancing different stakeholder needs. However, a secure and mediated VRR has greater potential for granting access in an ethical and equitable way than open access digital collections. VRRs have the potential to empower cultural heritage workers when engaging with communities and, in turn, empower communities to have better control of the records they steward, thus preserving relationships with the communities who create and research archival collections.

Culturally Sensitive Materials

As a result of colonization, universities, historical societies, and other non-tribal organizations have been custodians of most collections documenting Native Americans (O'Neal 2015). These take the form of manuscripts and recordings and were acquired without consent from tribal communities. As part of a larger movement towards self-determination and sovereignty, the development of tribal archives within these communities counters the destruction of culture and collective memory caused by

conquest. Tribal archives respect Indigenous epistemology and collection management (O’Neal 2015). However, issues remain surrounding the care, management, and protection of collections at non-tribal repositories.

Large-scale digitization and the creation of open access collections without engagement with tribal communities risks the violation of cultural protocols, exposing sacred and traditional knowledge. Protocols are specific to tribes and determine access to knowledge based on criteria such as age, gender, kinship, or seasonal activities (Christen 2015). Thus, the practice of making collections freely available for use—often in the public domain or under Creative Commons licenses—using common content management systems can be harmful to tribal communities. In contrast, VRRs can offer an environment for secure, mediated access to materials requiring different access levels. This approach to access and use is akin to the features of Mukurtu, an open source content management system designed for tribal communities to manage and share digital cultural heritage (Mukurtu 2021). Mukurtu enables communities to set their own levels of access based on cultural protocols. VRRs can also offer this flexibility by providing access to select users, therefore ensuring Indigenous values and ways of knowing are respected and knowledge is not inappropriately circulated.

In addition to providing the appropriate level of access to culturally sensitive materials, VRRs can also support repatriation efforts. Although physical materials have been legally owned by non-tribal institutions, often geographically distant from source communities, some institutions have repatriated objects to their tribal custodians. In consultation with the tribal community, a VRR could also be used to provide remote, mediated access to digitized versions of the repatriated objects. However, when physical return is not an immediate option because of insufficient storage facilities or internal politics, digital repatriation can be an alternative form of returning cultural heritage to the originating community (Bell, Christen, and Turin 2013). Although digital surrogates are not a substitute for the original materials, digital repatriation is one step towards restorative justice for tribal communities. VRRs can serve as a platform when selecting materials and negotiating digital repatriation projects with communities at a distance.

In tandem with building trusting relationships with tribal communities, VRRs can support institutions in the culturally responsive stewardship of digitized cultural heritage. For example, between 2018-2020 UC Irvine Special Collections & Archives managed an IMLS grant that sought to investigate how academic archival repositories can sustainably partner with community archives vis-a-vis a post custodial model to maximize access to their holdings (Eagle Yun et al, 2020). Resulting work included the development of a web-based form for community archives partners to ingest their materials directly into UC Irvine’s DAMS, Nuxeo. This agreement allows independent community partners to sustainably ingest their own content on Nuxeo but falls short on providing mediated access to end users. Future opportunities to extend this work include integration with a VRR system that will allow community partners to set mediated access to materials they do not wish to make openly available on the web.

Donor, Creator and Subject Privacy

Open, online access can threaten donor and creator privacy, consequently posing risks in our relationships with these groups and for the institution. This is especially true when digitizing collections which were donated before mass digitization became a common practice (Robertson 2018) and for collections where it may be unclear whether records are protected by federal privacy legislation. For example, the Health Insurance Portability and Accountability Act (HIPAA) protects medical records for

50 years after a patient's death. However, a repository may not know if the individual is deceased. Additionally, records with patient information may have been donated by a noncovered entity, an institution not subject to HIPAA (Gilliland and Wiener 2011). There is also ambiguity around the Family Educational Rights and Privacy Act (FERPA) and what constitutes a protected education record (Chute and Swain 2004). Other privacy needs to consider belong to third parties or subjects represented in archival materials. Images of deceased ancestors and some materials documenting historical or contemporary trauma should not be published online out of respect to affected communities.

Mediated access in a secure VRR environment reduces risk to affected communities and the institution. VRRs offer flexibility for collections with known restrictions. For example, donors may request that materials not be made widely available online but may allow physical and virtual reading room use. The institution can demonstrate it has made a good faith effort to protect the privacy of donors, creators, and third parties by providing a Terms of Use agreement stipulating that the researcher is responsible for not disclosing or publishing any materials. The VRR could require authentication and potentially block the copying or downloading of materials when deemed necessary for collection restrictions. In addition to federal privacy laws, archivists should also understand state laws pertaining to electronic records and health and medical information.

Recommendations: Ethical Concerns

- *Understand the scope of collections which represent tribal communities, and proactively build relationships with them*
- *Make sure community members can access the VRR, and collaborate with them to determine culturally appropriate levels of access*
- *Consider physical or digital repatriation when negotiating VRR access*
- *Understand what collection materials are subject to federal and state privacy legislation*
- *Consider drafting gift agreements that enable donors to waive their HIPAA and FERPA protected materials*
- *Implement proper safeguards to protect privacy and restricted materials within the VRR*
- *Provide terms and conditions of use (approved by general counsel) to VRR users*

Copyright

Copyright law affects the work of libraries and archives in pervasive and complex ways. As libraries and archives digitize collection items and collect more born-digital items that are protected by copyright, copyright is often cited as a constraint to expanding broad access to these collections.

The scale and accelerated pace of digitization and collection of born-digital works has exposed an additional issue. Libraries and archives do not have the resources to always make item-level determinations about the copyright and reuse status of each item. In an archival context, these determinations may prove impossible to make because the items are not described at the item-level.

Therefore, many determine that because there may be some works protected by copyright, none of the items in a folder, container or collection can be made available openly.

Although copyright law grants rights holders certain exclusive rights, those rights are tempered and balanced by a set of exceptions and limitations. These exceptions and limitations are essential to accomplishing the mission of libraries and archives. In situations where the rights status of collection items is not clear or when the risk of making items broadly available is too high for an institution, the exceptions and limitations in the law permit access to items in limited ways such as a VRR.

Exceptions, Limitations and VRRs

Libraries and archives enjoy a number of exceptions and limitations in copyright law that help them accomplish their mission. These exceptions and limitations are designed by Congress to permit limited uses of items without the permission of the rightsholder. Exceptions and limitations safeguard fundamental individual user rights, including the freedom of access to information and the freedom of speech.

The exceptions and limitations under the copyright law of the United States most relevant to the implementation of VRRs are the libraries and archives exceptions (17 U.S.C. §108), the right of fair use (17 U.S.C. §107), and the first sale doctrine (17 U.S.C. §109). Taken together, these exceptions provide the legal basis for VRRs.

The libraries and archives exceptions in Section 108 permit libraries to make certain uses of works protected by copyright to advance the goals and objectives of libraries and archives. Those uses include making copies for preservation and replacement purposes and making copies in response to patron requests. Section 108 was created in recognition of the special place libraries and archives hold in the production of knowledge.

In Section 108, sections (d) and (e) are the most relevant to VRRs. These sections permit the copying of collection items at the request of patrons solely for private study, scholarship or research. For example, libraries are permitted to make a copy of a small part of a book protected by copyright at the request of a patron. The right to make these copies, however, is not unlimited—there are specific requirements that libraries and archives must comply with to take advantage of this right. This right is also limited mostly to textual items; musical, visual and audiovisual works are excluded.

Another exception essential to cultural heritage organizations is the right of fair use. Unlike the prescriptive exceptions found in Section 108, fair use is a broad exception that permits a wide variety of uses without needing the permission of rights holders. Courts rely on a four-factor test to determine whether a particular use is permitted by fair use.

One of these factors is the impact of a use on the market for the work protected by copyright. For example, a magazine publishing the most important excerpts of a soon-to-be-published book would harm the market for that book (471 U.S. 539, 1985). When designing VRRs, developers will need to consider how their design will impact the market for the original work. Taking steps to limit the reproduction of collection items in the VRR will help enhance their argument that the VRR is providing only access to the original work, not permitting unlimited and unrestricted copying that may harm the market for the original. There are a number of tools that can be used to limit copying that might be

considered by VRR developers: DRM, secure streaming, and watermarking all help to discourage copying.

Another fair use factor courts use to analyze whether a particular use is permitted is the first factor which looks at the purpose and nature of the use. For example, providing digital reproductions of books to enable readers with print disabilities to access the knowledge contained in those books has been determined to be a fair use (755 F.3d 87, 102 (2d Cir, 2014). Educational and non-profit uses are viewed favorably by courts on this first factor. For VRRs, most are designed to support remote access to collection items to advance research and education. VRR developers can enhance their argument that their VRR is a fair use by considering how to actively limit the use of VRRs to patrons who are engaged in “private study, scholarship or research,” a use already viewed favorably by Congress. Some VRRs require users to agree to terms that explicitly limit the use of items found in VRRs to uses that are more likely to be considered fair use, while other VRRs ask patrons to provide more information about their interest in collection items to screen for uses that advance research.

The right of first sale is another important exception for VRRs. This right permits the rightsholder to control the first distribution of a particular copy of their work, but not subsequent distributions of that copy. This right is essential to libraries because it permits lending of collection items without requiring libraries to pay a license fee each time a book is borrowed. In the context of VRRs, this right of first sale may provide additional legal support. In particular, the concept of “Controlled Digital Lending” may provide a legal foundation to permit the limited access of digitized copies of collection items (Hansen 2018). Although VRR implementations and Controlled Digital Lending may share some important distinctions, they share the idea that libraries should be able to provide access to their collections for which they have a first sale right if reasonable steps are taken to limit the market harm to rightsholders.²

Combined, these exceptions and limitations spell out a clear desire by Congress to advance the progress of knowledge by permitting limited copying and distribution for the purposes of private study, scholarship or research. VRRs designers can strengthen their arguments that their VRRs fit squarely into these exceptions and limitations by taking steps to ensure that VRRs support the limited goals of providing remote access to collection items for limited purposes.

Recommendations: Copyright

- *Document legal rationales that underlie decision to provide potentially in-copyright collection items to remote patrons*
- *Consider taking steps to limit the reproduction of collection items that are protected by copyright in the VRR to enhance fair use arguments*
- *Consider developing parameters that explicitly define or limit the use of in-copyright collection items made accessible through the VRR for “private study, scholarship or research” (e.g., user agreements that limit uses, requiring patrons to submit information about their research interests, conducting research interviews with patrons seeking VRR access, etc.)*

² Controlled Digital Lending is the subject of a copyright lawsuit today. Depending on the resolution of that lawsuit, we may learn more about the applicability of the concepts of Controlled Digital Lending to VRRs.

Systems Interoperability

Meaningful holistic development of VRRs must exist in an ecosystem with other archival management, discovery, and delivery systems such as catalogs, finding aids, digital repositories, digital asset managers, and request management software. A significant challenge for VRR development has much less to do with the technology necessary to render a digital object on a user's computer and more to do with integrated management of existing systems and methods to make those objects (and information about these) discoverable and deliverable. None of the VRR systems represented in Appendix A represent a perfect interoperability scenario. Our VRRs range from standalone systems (UCI) to some level of interoperable compatibility with request management software (UCSD). An ideal interface for a VRR must deliver digital collections, digitized or born-digital, to a user; but is also seamlessly integrated with other archival management, discovery, and delivery systems in the back and front-end. It must do so while respecting the considerations in previous sections. If those considerations are met, the delivery of those files may appear very similar to remote, unrestricted access.

Existing initiatives and platforms such as the ePADD, Emulation as a Service Initiative (EaaSI), and the International Image Interoperability Framework (IIIF) have the potential to influence development and integration of archival discovery and delivery systems and help set a technological precedent for future development of VRRs. These initiatives provide examples of how authentication can be added into the provision of remote access to digital collections.

ePADD includes two separate modules for restricted and unrestricted access to email archives (Stanford University Library 2021). The Discovery module provides access only to the redacted versions of emails. The Delivery module provides full access. In the current recommended deployment, Discovery is used as a remote access tool for a user to identify potential material of interest, while Delivery is only accessible on-site. However, both modules use the same interface and underlying technology.

EaaSI renders born-digital objects by launching emulation environments on remote servers (Yale University Library 2021). Users access an environment through a browser window using remote desktop technologies, in effect working with a screencast of the environment rather than the computer itself. This prevents a user from being able to download files for use outside of the environment unless they are allowed by the application.

The IIIF APIs to publish and access digital still images, audio, video, and 3D scans include one for Authentication (Appleby et al., eds., 2017). The API does not implement a specific authentication standard itself. Instead it provides a method to interact with authentication protocols already in-place at the publishing institution. Once a request is authenticated, the API delivers the requested file or portion of the file.

None of these platforms require an authentication protocol, but they can all be implemented with one. In example deployments, users find objects that can be accessed through an existing discovery platform. In requesting access to the object, users may be presented with an authentication prompt, whether that is for a username and password, a clickthrough agreement to respect intellectual property rights, or instructions on how to access in a reading room.

The systems presented here are not the sole methods of creating remote access to collections. They are field-level initiatives that could be incorporated into a VRR. The specific requirements of the people represented within a collection, the formats of the material, the needs of the users, and resources of the organization necessitate a flexible approach to the technical platforms that make up a VRR. For example, North Carolina State University Libraries is configuring a virtualized environment through its campus Virtual Computing Lab for use as a VRR.

Recommendations: Systems Interoperability

- *Future VRR development must prioritize meaningful front- and back-end integration with other archival management, discovery, and delivery systems such as catalogs, finding aids, digital repositories, digital asset managers, and request management software*
- *VRR communities of practice must pay attention or collaborate with initiatives which have the potential to set technological precedents and user expectations for digital collection access like ePADD, Emulation as a Service Initiative (EaaSI), and the International Image Interoperability Framework (IIIF)*

Opportunities Moving Forward

VRRs have the opportunity to play a key role in the stewardship and access of archival collections. The authors hope the **Recommendations** listed at the end of each section (a chart with abbreviated recommendations can be found in Appendix C) provide a foundational, but in no way definitive, list of best practices towards responsible stewardship of VRR mediated digital discovery and access to archival collections.

VRR systems, practices, and services are far from operational maturity in the cultural heritage profession. Sharing knowledge across professional groups, research and its dissemination, are all key for further development on shared recommendations and practices for VRRs. Opportunities for further exploration include:

- building cross-institutional partnerships to collaborate on system technical requirements development and service best practices
- tapping into local, state-wide, national (international may be difficult due to varying copyright laws) working groups or existing library consortia structures to develop VRR communities of practice
- contributing to research on existing VRR models by developing survey opportunities (similar to Appendix A). Surveys may place emphasis on technological requirements; user/usability; copyright interpretation, alignment with sustainability work, etc.
- exploration of system integration and interoperability with other essential systems for archival access and discoverability
- exploration of creating a mailing list or group where sharing ideas and discussion can occur

Each of these opportunities can open the door to not only developing VRRs but enable libraries, archives, and museums to increase engagement with collections, address the needs of users in virtual settings—all while centering human-first priorities like staffing sustainability, cultural sensitivity and equitable access.

References

- Appleby, Michael, Tom Crane, Robert Sanderson, Jon Stroop, and Simeon Warner, eds. 2017. "Authentication." IIF Authentication API 1.0. Accessed on June 25, 2021. <https://iif.io/api/auth/1.0/>
- Authors Guild v. HathiTrust, 755 F.3d 87, 102 (2d Cir. 2014).
- Bell, Joshua A., Kimberly Christen, and Mark Turin. 2013. "Introduction: After the Return." *Museum Anthropology Review* 7 (1-2): 1-21. <https://scholarworks.iu.edu/journals/index.php/mar/article/view/3184>.
- Blumenthal, Karl, Peggy Griesinger, Julia Y. Kim, Shira Peltzman, and Vicky Steeves. 2020. "What's Wrong with Digital Stewardship: Evaluating the Organization of Digital Preservation Programs from Practitioners' Perspectives," *Journal of Contemporary Archival Studies* 7, Article 13: 1-23. <https://elischolar.library.yale.edu/jcas/vol7/iss1/13>
- Carnegie Hall. n.d. "Explore Watch and Listen." Accessed June 25, 2021. <https://www.carnegiehall.org/Explore/Watch-and-Listen>
- Christen, Kimberly. 2015. "Tribal Archives, Traditional Knowledge, and Local Contexts: Why the 's' Matters." *Journal of Western Archives* 6 (1): 1-19. <https://doi.org/10.26077/78d5-47cf>.
- Chute, Tamar, and Ellen Swain. 2004. "Navigating Ambiguous Waters: Providing Access to Student Records in the University Archives." *The American Archivist* 67 (2): 212-233. <https://doi.org/10.17723/aarc.67.2.v418724687421m15>.
- Eagle Yun, Audra, Jimmy Zavala, Krystal Tribbett, and Thuy Vo Dang. 2020. "Transforming Knowledge, Transforming Libraries - Researching the Intersections of Ethnic Studies and Community Archives: Final Report." *UC Irvine: Libraries*. <https://escholarship.org/uc/item/47c2h0dd>
- European Union. 2018. "General Data Protection Regulation (GDPR)." Accessed June 25, 2021. <https://gdpr-info.eu/>.
- Farmer, A. 2018. "Archiving While Black." *The Chronicle of Higher Education*. Accessed June 15, 2021. https://www.chronicle.com/article/archiving-while-black/?bc_nonce=s0adbcp463pobxswgid8zl&cid=reg_wall_signup
- First Archivists Circle. 2007. "Protocols for Native American Archival Materials." Accessed June 25, 2021. <https://www2.nau.edu/libnap-p/>
- Gilliland, Anne T., and Judith A. Wiener. 2011. "Digitizing and Providing Access to Privacy-Sensitive Historical Medical Resources: A Legal and Ethical Overview." *Journal of Electronic Resources in Medical Libraries* 8 (4): 382-403. <https://doi.org/10.1080/15424065.2011.626347>.

- Hansen, David and Kyle Courtney. 2018. "A White Paper on Controlled Digital Lending of Library Books." Accessed on June 28, 2021. <https://controlleddigitallending.org/whitepaper>.
- Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 539 (1985).
- Light, Michelle. 2014. "Managing Risk with a Virtual Reading Room: Two Born Digital Projects." In *Reference and Access Innovative Practices for Archives and Special Collections*, 17-35. Lanham, MD: Rowman & Littlefield Publishers. https://digitalscholarship.unlv.edu/lib_articles/463/
- Matienzo, Mark A., Dinah Handel, Josh Schneider, and Camille Villa. 2020. "Lighting the Way: A Preliminary Report on the National Forum on Archival Discovery and Delivery." Stanford University Libraries. Accessed June 25, 2021. <https://purl.stanford.edu/rt078dm2196>
- Mukurtu CMS. n.d. "Home." last accessed June 24, 2021, <https://mukurtu.org/>.
- New York Public Library. 2020. "Welcome to the Virtual Reading Room: Digitized Archives for Home Research." last modified May 20, 2020. <https://www.nypl.org/blog/2020/05/04/welcome-virtual-reading-room>
- O'Neal, Jennifer R. 2015. "'The Right to Know': Decolonizing Native American Archives." *Journal of Western Archives* 6 (1): 1-17. <https://doi.org/10.26077/fc99-b022>.
- Princeton University Library. n.d. "Virtual Services." Accessed June 25, 2021. <https://library.princeton.edu/special-collections/services/virtual-services>
- Princeton University. 2021. "Web Accessibility Guidelines." last modified April 9, 2007. <https://accessibility.princeton.edu/guidelines/web>
- Research Libraries UK, University of Manchester. 2021. "Virtual reading rooms speed presentation." May 4, 2021. <https://www.youtube.com/watch?v=TWKRRyrgV0&t=1s>.
- Robertson, Tara. 2018. "Not All Information Wants to be Free: The Case Study of On Our Backs." In *Applying Library Values to Emerging Technology: Decision-Making in the Age of Open Access, Maker Spaces, and the Ever-Changing Library*, edited by Peter D. Fernandez and Kelly Tilton, 225-239. Chicago: Association of College and Research Libraries. <http://eprints.rclis.org/32463/>
- Selinger, Evan. 2013. "Technological Fixes." *The Wilson Quarterly* 37 (2). <https://www.jstor.org/stable/wilsonq.37.2.11>.
- Society of American Archivists Electronic Records Section. 2016. "Born-Digital and in the Virtual Reading Room." February 11. <https://saaers.wordpress.com/2016/02/11/born-digital-and-in-the-virtual-reading-room/>
- Society of American Archivists. n.d. "Virtual Reading Room." Accessed June 25, 2021. <https://dictionary.archivists.org/entry/virtual-reading-room.html>.
- Society of American Archivists. 2019. "Guidelines for Accessible Archives for People with Disabilities." Accessed June 25, 2021. https://www2.archivists.org/sites/all/files/SAA%20Guidelines%20for%20Accessible%20Archives%20for%20People%20with%20Disabilities_2019_0.pdf

- Stanford University Libraries. 2019. "Lighting the Way: illuminating the future of discovery and delivery of archives." April 19, 2019. <https://library.stanford.edu/blogs/stanford-libraries-blog/2019/08/lighting-way-illuminating-future-discovery-and-delivery>
- Stanford University Libraries. 2021. "ePADD." Accessed June 25, 2021. <https://library.stanford.edu/projects/epadd/>.
- Tactical Tech, Mozilla. 2017. "The Data Detox Kit." Accessed June 25, 2021. <https://datadetoxkit.org/>.
- University of California Berkeley. 2021. "Types of Assistive Technology." Accessed on June 25, 2021. <https://webaccess.berkeley.edu/resources/assistive-technology>
- University of California Berkeley. 2021. "Types of Assistive Technology." Accessed June 15, 2021. <https://webaccess.berkeley.edu/resources/assistive-technology>
- University of California at Irvine Libraries. "Application for the Virtual Reading Room in UCIspace." Accessed June 25, 2021. <https://special.lib.uci.edu/application-virtual-reading-room-ucispace>
- Weber, Chela Scott. 2020. "Time Estimation for Processing Born-Digital Collections". *Hanging Together, the OCLC Research Blog*, April 28, 2020. <https://hangingtogether.org/?p=7911>
- Weber, Chela Scott, Martha O'Hara Conway, Nicholas Martin, Gioia Stevens, and Brigitte Kamsler. 2021. *Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections*. Dublin: OCLC Research. <https://doi.org/10.25333/zbh0-a044>.
- World Wide Web Consortium (W3C) Web Accessibility Initiative (WAI). 2021. "Web Content Accessibility Guidelines (WCAG)." Accessed June 25, 2021. <https://www.w3.org/WAI/standards-guidelines/wcag/>.
- Yale University Library. 2021. "Yale Library Emulation Viewer: Home." last modified May 18, 2021. <https://guides.library.yale.edu/emulation-view/>

Appendix A: Policies and Feature Commonalities

The following table illustrates policies, guiding principles and common features considered/implemented in VRRs for the organizations participating in this writeup. Column definitions are available in Appendix B.

Organization	System	Covered by Universal Agreement or Policy?	Personalized User Agreement / Interview Required?	Requires Authentication?	Allows Guest Users?	Allows Download?	Uses Watermarks?	Can Embargo?
Carnegie Hall	Cortex (public collections)	Yes	No	No	Yes	Yes	No	Yes
	Cortex (restricted collections)	Yes	Yes	No	Yes	Yes	No	Yes
NCSU	Virtual Computing Lab (VCL)	TBD	Yes	Yes	Yes	No	No	No
NYPL	N/A (not yet implemented)	TBD	Yes	Yes	Yes	No	Yes (encrypted streaming)	No
Princeton	Custom (Samvera) (open collections)	Yes	No	No	Yes	Yes	No	No
	Custom (Samvera) (restricted collections)	Yes	Yes	Yes	Yes	No	No	No
	Custom (Samvera) (public collections controlled lending)	Yes	No	Yes	Yes	No	No	No
San Bernardino County	Custom and Dropbox	Yes	Yes (user interview)	No	Yes	No	No	No
UC Irvine	DSpace (pre 2018), OneDrive and Google Drive	Yes	Yes	Yes	Yes	Yes	No	Yes
UCSD	Custom (Fedora-like)	Yes (Aeon)	No	Yes	Yes	No (Exception: PDF files)	Yes (for image and text files)	No

Figure 1: Policies Table, columns 1 - 8

SPEEDING TOWARDS REMOTE ACCESS: DEVELOPING SHARED RECOMMENDATIONS FOR VIRTUAL READING ROOMS

Organization	Limited Time for Access?	Limited Number of Users?	Online Access?	Off-Site Access?	Analytics Enabled or other forms of User Tracking?	Open Source Software?	ADA or WCAG Compliant?	Level of Institutional Support	Staffing?
Carnegie Hall	No	No	Yes	Yes	Yes	No	Audited; remediation ongoing	Core service	Dedicated staff
	Yes	No	Yes	Yes	Yes	No	Audited; remediation ongoing	Supplemental	Dedicated staff
NCSU	Yes	No	Yes	Yes	No	VCL is open source, but users connect using preferred software.	TBD	Pilot/ Experimental	Taken on as new duties
NYPL	No	No	Yes	Yes	Yes	No	N/A	Core service	Taken on as new duties
Princeton	No	No	Yes	Yes	Yes	Yes	Audited; remediation ongoing	Core service	Dedicated staff
	No	Yes (admin-specified per-collection)	Yes	Yes	Yes	Yes	Audited; remediation ongoing	Core service	Dedicated staff
	Yes	Yes	Yes	Yes	Yes	Yes	Audited; remediation ongoing	Core service	Dedicated staff
San Bernardino County	No	No	Yes	Yes	Yes	No	remediation ongoing	Core service	Dedicated staff
UC Irvine	Yes	No	Yes	Yes	Yes	Yes	remediation ongoing	Core service	Dedicated staff
UCSD	Yes	No	Yes	VRR-designated materials are accessible remotely, with no geographical or IP restrictions. We do have a few digital collections more generally, that are IP-restricted.	Registration and Request data maintained in Aeon System, and available for reporting purposes	Yes	N/A	Core service	Taken on as new duties

Figure 2: Policies Table, columns 9 - 17

Appendix B: Column Definitions

Column definitions for the table represented in Appendix A - Policies table (figures 1 and 2).

- **Organization**
 - Participating organization in the Lighting the Way Virtual Reading Rooms Group
- **System**
 - Software/platform used to power to power Virtual Reading Room functionality
- **Covered by Universal Agreement or Policy?**
 - Does access/use fall under broad or general policies or agreements for the institution?
Examples include: Privacy Policy, Terms and Conditions, Codes of Conduct.
- **Personalized User Agreement / Interview Required?**
 - Do users need to provide additional information and/or agree to additional terms (not part of Universal Agreement or Policy) to gain access?
- **Requires Authentication?**
 - Does access require user identity verification? Examples: Central Authentication Service (CAS), multi-factor authentication (MFA), confirm email address.
- **Allows Guest Users?**
 - Are there ways for users outside of the organization to access (for example, sponsored guest accounts)?
- **Allows Download?**
 - Can users download material from VRR?
- **Uses Watermarks?**
 - Are watermarks applied or optionally applicable to content accessible via VRR?
- **Can Embargo?**
 - Can the VRR restrict access on the object or collection level until a specified date?
- **Limited Time for Access?**
 - Does user access expire in a reasonable period? Expiration of access can be a range e.g., 1 hour to 1 year.
- **Limited Number of Users?**
 - Is there a cap on how many users are allowed access simultaneously?

- **Online Access?**
 - Is VRR accessible via a web browser or application to users (e.g., in a web browser, remote desktop)?
- **Off-Site Access?**
 - Is VRR accessible remotely, or otherwise restricted by geographical location or other factor (e.g., on-campus VPN access, IP restrictions)?
- **Analytics Enabled or other forms of User Tracking?**
 - Is user activity, behavior, and data being actively collected, monitored, and managed? This can include Google Analytics, other commercial analytic tools, or strategies utilizing data to report on users.
- **Open Source Software?**
 - Is the VRR solution using majority or entirely Open Source software?
- **ADA or WCAG Compliant?**
 - Is the VRR compliant with U.S.-based accessibility requirements? Currently, the Americans with Disabilities Act (ADA) does not define "web accessibility" requirements. The W3C's Web Content Accessibility Guidelines (WCAG) identify 3 levels of conformance [A (bare minimum level of accessibility), AA (target level of accessibility meeting legal requirements) and AAA (exceeds accessibility requirements)] for 4 Principles (Perceivable; Operable; Understandable; Robust).
- **Level of Institutional Support**
 - The level at which VRR work is currently prioritized at an organizational level. Factors may include: funding, funding type, staffing, resources like access to legal expertise or subject matter experts. "Ad-hoc/Responsive" can refer to servicing requests through file sharing services like Dropbox, email attachments, etc. "Interest but not resources" can refer to an institution that has a strategic or functional need but cannot prioritize or plan for a solution at this time.
- **Staffing?**
 - The level at which VRR efforts are currently staffed within an organization. "Taken on as new duties" can refer to supplemental staff or responsibilities that are critical to the VRR functioning, but are not in scope of dedicated/resourced position responsibilities. "Precarious staffing" can refer to precariously funded or time-limited positions, paid or volunteer, including interns, fellows, and student workers.

Appendix C: Recommendations

An abbreviated list of all recommendations made in the report.

Advocacy and Outreach	Resources	Users and Use of VRRs	Ethical Concerns	Copyright	Systems Interoperability
Gain administrative support by emphasizing the legal basis for VRR, and popular benefits like secure equitable expanded access, cutting edge technology, and culturally sensitive access methods. Build case around institutional strategic priorities (e.g. improve discoverability and use of archival materials)	Establish permanent funding source for VRRs before development of the system and service	Establish user access requirements, listing needs for various groups of users. Program into platform access based on VRR policies; including authentication of users, limitation based on location, and click through agreements	Implement proper safeguards to protect privacy and restricted materials within the VRR. Provide terms and conditions of use (approved by general counsel) to VRR users	Document legal rationales that underlie decision to provide potentially in-copyright collection items to remote patrons	Future VRR development must prioritize meaningful front- and back-end integration with other archival management, discovery, and delivery systems such as catalogs, finding aids, digital repositories, digital asset managers, and request management software
Be prepared to articulate the importance of granular access options and relevant workflows and practices for digital archives when speaking with external systems developers, administrators, general counsel, users, donors, and other key stakeholders	Developing and sustaining a VRR system will require some level of in-house software development expertise - ensure adequate permanent staff to maintain and develop VRR; consider adding language to job descriptions or annual plans specific to VRR development and management	Establish policies and protocols to secure user data collected for using the VRR. Establish protocols for mediating access, such as time based access and disabling downloading capabilities	Understand what materials are subject to federal and state privacy legislation. Consider drafting gift agreements (approved by general counsel) that enable donors to waive their HIPAA and FERPA protected materials	Consider taking steps to limit the reproduction of collection items that are protected by copyright in the VRR to enhance fair use arguments	VRR communities of practice must collaborate with initiatives like ePADD, Emulation as a Service Initiative (EaaS), and the International Image Interoperability Framework (IIIF) which have the potential to set technological precedents and user expectations for digital collection access

Figure 3: Recommendations Table, rows 1 - 3

SPEEDING TOWARDS REMOTE ACCESS: DEVELOPING SHARED RECOMMENDATIONS FOR VIRTUAL READING ROOMS

Advocacy and Outreach	Resources	Users and Use of VRRs	Ethical Concerns	Copyright	Systems Interoperability
Engage the necessary stakeholders at the onset of developing VRR and include software developers, access services staff, technical services, records management team, research data unit, general counsel, users, and communities with culturally sensitive needs throughout the development process	Build in specific time to execute the development of the VRR and account for the professional development needs staff will need to initiate and maintain a successful service (skill building, knowledge sharing, research)	User testing should be done on a regular basis. Frequently engage with research communities to learn how they use VRR and how well it meets their needs. Use a low barrier feedback mechanism so that users can provide feedback. Evaluate gaps and modify the service to reach the most users. Create frequent training and marketing opportunities so that users understand the VRR and how and when to use it	For tribal collections, proactively build relationships with the community. Understand the scope and access needs of collections. Make sure community members can access the VRR, and collaborate with them to determine culturally appropriate levels of access	Consider developing parameters that explicitly define or limit the use of in-copyright collection items made accessible through the VRR for “private study, scholarship or research”	
	Ensure all staff are included in discussions and fact gathering for development of VRR including software developers, maintenance, public service, reference, instruction, digital archivists and curatorial staff	VRRs must work with assistive technology. Screen readers, text to speech apps, screen magnification software, and other assistive technologies should not be an afterthought in development. Many institutions have minimum web access accessibility requirements that can help guide project teams through development	Consider using VRR for physical or digital repatriation of tribal collections		
		There may already be an accessibility team that can help evaluate a new web service at your institution, system developers may be familiar with best practices, and users can tell your institution when something does not work for them			

Figure 3: Recommendations Table, rows 4 - 8