REMOTE ACCESS TO ARCHIVES AND SPECIAL COLLECTIONS

Written by Wes Hamrick,
Kimberly Kennedy, and Garrett McComas

sourcery

Northeastern University
Library

GREENHOUSE STUDIOS

Table of Contents

Executive Summary

Organized by members of the Sourcery team at Greenhouse Studios at the University of Connecticut and staff at the Northeastern University Library and supported by the Andrew W. Mellon Foundation, the conference series "Remote Access to Archives and Special Collections" took place in the fall of 2020. The Sourcery app provides a platform for delivering photographs or scans of documents in archival collections to remotely located researchers. The series used Sourcery to spark discussion of how to serve researchers accessing archives and special collections materials remotely. The organizers asked participants to consider new approaches to remote access and to rethink some of their current practices. One of the major themes to emerge from the sessions was the tension between practitioners' desire to build digital collections with high-quality images and detailed metadata, and many researchers' need for quick access, with quality a less important concern. The other major theme was the high volume of digitization requests archives receive in addition to their other important preservation, description, and teaching work. With budgets restricted at many institutions, creating sustainable digitization workflows can be difficult.

I. Broader Issues

While each of the five workshop sessions raised a number of specific issues concerning remote access, there were at least two points of discussion that were threaded throughout the workshops. These broader concerns suggest additional areas of research for the Sourcery project moving forward, as well as directions for future development of the platform.

Discussions of Reference Scans and Access

One of the major themes that was present across each of the sessions was the discussion of the production and use of reference scans. Here, the term "reference scan" refers to lower-quality images that can be quickly produced to fulfill reference requests in archives and special collections. In many institutions, these types of scans are separate from higher quality digitization efforts in the way that they are approached conceptually and in practice. Reference scans are often fulfilled by archivists, have little to no associated metadata, and are used to fulfill the information needs of a researcher. They are not meant for publication and normally not meant to to be stored for future use. High-quality digitization efforts are often conducted by specialists and have detailed metadata. They are meant to transcend simply fulfilling an information need, allowing researchers to more closely investigate the document through diverse means such as zooming in to look at fibers or text mining. They are meant to be stored indefinitely and, in some ways, serve as a replication of the original document.

Broadly speaking, high-quality digitization efforts have recently received resources, from grants and from institutions themselves, since they serve as a way to highlight the institution's collections. Less attention and fewer resources have been directed towards reference requests, although they are much more efficient at fulfilling the information needs of most potential patrons. This speaks to the importance of the conversations that occurred during the workshop series, as this area has been under-examined and often under-served.

Across the sessions there was discussion of how reference scans could be used in new and innovative ways. Many of the participants were interested in how Sourcery might be integrated with Tropy to create a crowd-sourced digital archive of document photographs taken by researchers. This was particularly attractive to those that work in archives and special collections, as it would not add additional labor requirements or disrupt current workflows, both of which are particularly relevant with the uncertainties resulting from the COVID-19 pandemic. Some institutions are already exploring something similar to this with their own collections, such as the Folger Shakespeare Library which provides online access to its reference scans using LUNA[1], a management system for lower-quality

---

[1] https://luna.folger.edu/luna/servlet/FOLGERCM1~6~6

images. This type of approach follows a "digitization prior to description"[2] model that requires placing concerns of access before preservation and metadata. There are, of course, trade-offs and valid concerns to prioritizing lower-quality scans,[3] but there was interest among institutions for this type of repository, especially if the required labor and upkeep were not prohibitive.

Building Sustainable Practices of New Initiatives

The COVID-19 pandemic caused institutions to pursue new initiatives, such as remote video consultation and a higher volume of reproduction fulfillment. The pressure to preserve access to archives and special collections while travel and in-person viewing are restricted has already changed archivists' day-to-day work, as financial restrictions have caused many of them to take up these new responsibilities themselves. As in-person activities gradually resume, there will be pressure to restore pre-pandemic activities, while maintaining the new practices that were in response to the pandemic. This could cause some confusion and strain on archivists, who are already having difficulty fulfilling the number of requests they are receiving. The capacity to find sustainable solutions to these problems will vary across institutions, and will be constrained by the resources available to each. Institutions are still unsure of how the current pandemic will affect their budgets in the future and how long the constraints will last.

A large part of building sustainable practices is the division and support of labor involved in maintaining remote access to archives and special collections. Many of the participants voiced concerns about the amount of remote access requests that they were receiving. Along with the current lack of student and other labor, some archivists were concerned that adding Sourcery to an already full queue of requests would exacerbate the strain on staff. One solution is that Sourcery could add a revenue stream to employ more staff, but this would be in tension with some institutions offering remote access for free. Another solution would be for institutions to shift their perspective on the fulfillment of remote reference requests. As has been stated, institutions are already concerned about the number of remote reference requests, and adding another platform for users to create requests could just add to the overflow of work. However, some institutions may choose to alter the way that they fulfill the requests, such as using mobile devices or rethinking imaging or metadata quality standards in order to fulfill more requests. Any changes to current workflows presents tradeoff scenarios, as many institutions have already had to explore due to budget cuts and the COVID-19 pandemic. Moments like now can also serve to reorient or strengthen priorities and practices of institutions. The integration of Sourcery into sustainable workflows will continue to be explored as the project moves forward.

---

[2] Dimunation, Mark. "Everything We See Hides Another: Coping With Hidden Collection in the 21st Century Library." In *Something's Gotta Give: Charleston Conference Proceedings, 2011*, edited by Bernhardt Beth R., Hinds Leah H., and Strauch Katina P., 21-29.

[3] For example, the "quality dilemma" that Paul Conway lays out in *Preservation in the Age of Google* (2010).

II. Summary of Conference Series

The "Remote Access to Archives and Special Collections" conference series covered issues in remote access from multiple perspectives and took place over five Fridays from October 16th to November 20th, 2020. The topics of each session were:

A. October 16th: "Introduction to Remote Access to Archives and Special Collections", Keynote address by Dan Cohen, organized by Tom Scheinfeldt and Patrick Yott

B. October 23rd: "Archives and Special Collections Staff Discussion", Keynote address by Jess Colati, organized by Kimberly Kennedy

C. October 30th: "Administration and Finance Staff Discussion", Keynote address by Greg Colati, organized by Garrett McComas

D. November 6th: "Scholars and Researchers Discussion", Panelists included Hilary Bogert-Winkler, Mills Kelly, and Grace Kuipers, organized by Wes Hamrick

E. November 20th: "Closing Session", Keynote address by Cliff Lynch

Each session included an organizer and was headlined by a speaker or group of panelists. The organizer of each session created the format and agenda, including coordinating speaker(s), and gave an overview of the session at the beginning. The speaker for each session gave opening remarks to frame the conversation and highlight themes for that particular session. Each session was attended by at least 40 participants from a wide range of collecting institutions.[4]

As a part of the Sourcery project funded by the Andrew W. Mellon Foundation, this conference was envisioned as a series of workshops where members of the archives and special collections communities could communicate about current and future practices, opportunities, and challenges in regards to remote access. The series often touched on the potential use of Sourcery as a way to fulfill remote requests for archives and special collections and partly acted as community outreach for the Sourcery project. The discussions elicited a mix of interest, skepticism, misunderstanding, and constructive feedback with regard to Sourcery and provided crucial community input that will help shape the project's future development. As with everything during this time period, the ongoing COVID-19 crisis was a substantial part of the discussions, as it affected institutions in myriad ways.

---

[4] See Appendices A, B.

A. Opening Session

The opening session took place on October 16th, 2020. The keynote speaker was Dan Cohen, Vice Provost/Dean of the Library and Professor of History at Northeastern University. The facilitator was Tom Scheinfeldt, Director of Greenhouse Studios and Associate Professor of Digital Culture, Learning, and Advocacy at the University of Connecticut. Patrick Yott, the Associate Dean for Digital Strategies and Services at Northeastern welcomed participants to the series.

Cohen discussed how technology, particularly the digital camera, has changed archival research. With the widespread adoption of smartphones, the world now has billions of high quality cameras. This has had a huge effect on archival research. One study showed that 90% of researchers in archives took photos on their phone as part of their research strategy, and nearly 40% of those researchers took more than 2000 photos.[5] Archives and libraries are still trying to catch up with this massive change in research practices.

In archives, there have been mass movements to digitize archival collections to make them accessible remotely. This work is more important than ever during COVID-19, when travel is limited. Cohen also encouraged the group to think about how Sourcery might integrate with other digital platforms such as Tropy, a tool researchers use to catalog photos taken in archives. Archivists and librarians need to consider how institutions might adapt and build a system that is more responsive to the way researchers work.

After Cohen's remarks, Scheinfeldt gave a presentation on Sourcery (https://sourceryapp.org/). The Sourcery App was originally built to take advantage of individual smartphone capabilities by creating a platform where researchers could place requests that could be filled by other people near an archive or special collection. However, in response to COVID-19 and restricted travel, the Sourcery App has altered its focus. An institutional (or enterprise) version of Sourcery aims to provide users and archives with a simple and clean portal for placing requests and receiving archival scans.

Finally, there was a general discussion and question and answer period. There were many questions about Sourcery integrations. Scheinfeldt responded that the team is investigating integration with Tropy, as well as integrations with Zotero, Omeka, Aeon, and ArchivesSpace. More information about the future of the Sourcery app can be found in "Future Directions for Sourcery" in Section III of this paper.

Other participants raised questions about how Sourcery will affect traffic in a reading room, and what the ethics of a "gig economy" model for this work are. Cohen responded

---

[5] Ian Milligan, "Becoming a Desk(top) Profession: Digital Photography and the Changing Landscape of Archival Research," (Google slides, AHA 2020).

that institutions will need to figure out the balance between Sourcerers and the work of long-term staff. Cohen and Scheinfeldt also pointed out that this kind of work is already occurring informally. Professors often send grad students into an archive, or post on a message board asking someone to take a photo of a document for them. Sourcery offers a way of formalizing this process to provide access to more material in a better way. Two Greenhouse Studios members, Wes Hamrick, Fellow, and Brooke Gemmell, Design Technologist, previously interviewed graduate students already engaged in this work in 2019. Many are doing this, sometimes for no pay, as a favor to senior faculty members, and were excited about the approach Sourcery was taking. They want to be paid and have the flexibility to fit work around their teaching and coursework responsibilities. Other participants mentioned that there are communities of people who do independent archival research as their full-time job, and it would be helpful for the Sourcery team to get their perspective and address their potential concerns. The Sourcery team is mindful of these critiques, and wants to support and supplement archival work, not undermine it. The Sourcery team is currently working more closely with archives and special collections professionals to better understand how to integrate into current workflows and will look to reach out to independent researchers in the near future.

B. Archivists/Special Collections

The Archives and Special Collections staff session took place on October 23rd, 2020. The keynote speaker was Jessica Colati, a lecturer of History and Director of the Archives Track, University of Massachusetts, Boston. The facilitator was Kimberly Kennedy, Digital Production Coordinator at the Northeastern University Library. After Colati's remarks, participants broke into four breakout rooms for discussion, then returned to the main room for a question and answer period.

Summary of Jessica Colati's remarks

Colati began by stating that, though the core practices of archivists have not changed, during COVID-19 they are navigating a difficult situation with limited access to materials. Patrons can't be welcomed into exhibitions and reading rooms like usual, and issues around limited access and new processes of handling materials have arisen. For example, there are limitations on how many people can be in an office.[6] Archivists have to think about a contagious virus living on collections documents.[7] These safety considerations mean that archivists may not be working as closely with physical collections as usual.

There are also many challenges associated with the use of materials in curricula and research. Instructors have voiced concerns about the difficulty of keeping students

---

[6] CNI, "What Happens to the Continuity and Future of the Research Enterprise: Looking to 2020-2021 and Beyond," 2020.
[7] OCLC, "REALM project research," February 11, 2021.

engaged in courses in which they would normally work with archival materials directly. In addition, faculty members and graduate students are experiencing obstacles to completing their research. For many researchers in the humanities and other disciplines, libraries and archives are akin to labs, and they need access just like scientists need access to their labs.

However, the challenge of remote access to archives and special collections is not new. Archivists were well-versed in working and providing reference services remotely, and were able to switch quickly into providing primarily remote service. Archivists need to continue to develop new strategies and tools for providing access to collections. Archivists can't provide access the way they always have.

Summary of Discussions

The main theme that emerged during the discussions was the difficulty of managing an increased volume of requests amid expectations of quick service. This was challenging when staff were only allowed onsite on a limited basis. In addition, some archives rely heavily on student workers for reference digitization requests and lost their labor due to occupancy restrictions and budget cuts, something that participants suspected may persist after the end of the pandemic. Some institutions were entirely closed for an extended period of time, resulting in a backlog of requests. With these obstacles, it can be difficult to maintain current standards and keep up with the volume of requests. Another issue with reading rooms being closed or on reduced hours is that researchers aren't able to investigate for themselves, putting a greater burden on archivists to do research in preparation for reproduction requests.

As a way of dealing with this, some institutions are limiting the number of pages that can be requested or charging for requests. Conversely, others were waiving page limits and fees for scans in light of the pandemic. Many institutions had already removed fees for digitization, and there is a sense among some that a cost-recovery model isn't worth the trouble and isn't equitable.

Some institutions are also looking at how to prioritize requests from different groups of people, asking who the institution should serve first. One participant mentioned a tension often apparent in archives connected to universities. Many universities focus on providing services to their own students, faculty and staff first, but the audience for archives and special collections might be primarily off-campus. Some archivists find it difficult to serve those outside of their university system when resources are directed primarily towards those on campus or within the system. This has been a major issue and tension in the archives and special collections of research universities, who feel torn between devoting resources and staff time to internal and external audiences.

Another major topic was the quality of reproductions and how digitization might become more flexible to increase access. Standardized digitization is a time consuming process that requires special equipment as well as the creation of detailed metadata, but some wondered how archives and special collections could better leverage quick, low quality scans. Some are ramping up digitization with additional equipment, such as copy standards for high-resolution images created in the TIFF format. A downstream question is how these materials will be included in the institution's digital collections. Those processes typically require the creation of metadata, which can be labor intensive and expensive, resulting in a backlog in making digitized materials available on the web.

Institutions are also providing digital reference services that don't involve reproduction requests. Some archives talked about providing text-based chat service or live video calls. Others are creating videos of material for patrons. This can give a better sense of the materiality of collections than a flat scan.

General discussion

A theme central to this session's discussion was labor, which the pandemic has affected greatly. Work typically done by students or interns is now done by archivists, who are the only ones allowed in the buildings.[8] This increases the pressure on their time and changes normal workflows. While Sourcery cannot solve the problem of insufficient staff to provide reference scans for researchers, discussions around Sourcery have nonetheless surfaced a gap between researcher demand for remote access and the available resources to meet this demand. While our ultimate goal is to produce a tool that benefits both researchers and archival institutions, we are also hopeful that the questions and provocations Sourcery raises can help lead to solutions to long-standing problems, including those of archival labor.

There were several concerns raised about Sourcery. One was that archives already receive requests through multiple channels, some of which are duplicate requests. Sourcery could add another place to check, increasing the workload on archivists. A counterpoint is that duplication of requests is a function of researchers not knowing whether a message is going to get through. A single interface, like Sourcery, would be a better experience for the researcher and, if implemented widely across institutions, could reduce researcher confusion and thereby the number of duplicate requests.

One comment concerned the question of charging for scans. This is an institution-level decision, not a platform decision. Some institutions don't have a way to deal with the infrastructure of charging. They have determined that however much you charge for a scan will not be enough to cover costs of the labor of pulling boxes, researching, and potentially cataloging and adding items to the institutional repository. The Sourcery team

---

[8] By the distribution of this paper in late July, 2021, this may have changed in a number of institutions.

commented that in the enterprise version of Sourcery, institutions will be able to set pricing, and Sourcery will collect fees and process credit cards.

Participants also wondered how reference interviews will work within Sourcery. The Sourcery team acknowledged that they don't know yet. In Sourcery's conception, the idea was that a researcher would already have a specific citation, which could be from a previous visit, or from the footnotes of published papers by other scholars. In many cases, a research strategy is built around citations. However, several archivists expressed concern that historians' footnotes are often not complete enough to accurately locate the item in the archives. Despite the difficulties of incomplete citations, Sourcery could offer value to researchers that want to check out one document in an archive. Even before COVID-19, they would not have made a trip for one item, so Sourcery can fill in a gap.

The group also discussed the modern practice of researchers taking many photos in the archives. These are typically kept for personal use and not described or tagged, and could be considered a suboptimal use of resources. One participant wondered whether researchers can square their informal practices with the more formal practices of the institution. Should digitization by the archives and digitization by researchers remain entirely separate? This opens up technical questions, since library systems are often difficult programs with which to build integrations. Some suggested that Sourcery might integrate well with Tropy, a tool for managing digital content. This could have benefits for the researchers. Luckily, the existing Mellon funding for Sourcery does include planning for integration with Tropy and Zotero. One participant suggested that if a historian could make pieces of citable material available publicly, that could solve the problem without the need to involve archive and library systems.

Issues surrounding access to audiovisual and born-digital material archival were briefly discussed, but more research and discussion is needed. For example, some born-digital collections present privacy and security concerns. If typical practice had been to limit use to the reading room, how does that transfer to a remote environment? Audiovisual material can also present difficulties with copyright and large file size. Some solutions mentioned were negotiating access terms, using a course reserves system, using controlled digital lending, and sharing files through a customized FileNet system.

In conclusion, many in the group urged that archivists need to be more involved in development if Sourcery is going to be part of the reference queue. There were questions about plans to sustain these conversations, and there was a clear interest in continuing the conversation so that together we can find a system that works for all parties.

C. Administration/Finance

The Administration and Finance session took place on October 30th, 2020. The speaker was Greg Colati, the director of the Connecticut Digital Archive (CTDA) and the Digital Preservation Repository Program Director at the University of Connecticut. The organizer for the session was Garrett McComas, a Post-Doc Fellow at Greenhouse Studios, University of Connecticut. After Colati's remarks, we broke into four breakout rooms for discussion, then returned to the main room for a general discussion and question and answer period.

Summary of Keynote Speech

The aim of Greg Colati's remarks was to probe the financial and administrative position of archives and special collections within the larger library system. The differences between archives and special collections and the rest of a library system manifest in many ways. Preservation, access, and inventory management is different for unique collections as opposed to journals and monographic materials. At its most basic point, there are historically higher barriers to entry for unique materials than for mass produced materials. Archives and special collections are also often separate from the rest of a library system through administrative and budgetary framework. Perhaps, as Rossman argues, if general and special collections continue to be seen as separate entities instead of a "continuum of research resources"[9], it will negatively affect both researchers and special collections units. Colati questioned whether the current framework supports the work that archivists want to do and whether there might be ways to change that framework for the better.

Decisions about library holdings and materials are often made by assessing metrics taken from those collections. For example, some archives assess their value by the number of international scholars that access their collections. However, the current COVID-19 pandemic has caused people in many professions to rethink what they value and how they value it. In recent years, a lot of value has been placed on the digitization of collections, and all the labor and investment that comes with it. This includes digitization equipment, the creation of metadata, digital infrastructure (including websites), and the evaluation of collections metrics. Larger institutions are able to "outsource" some of this labor to cataloging departments and digitization specialists in other parts of the library system. Archivists, then, are more free to build outreach and education programs and argue for the value of their collections. Workers in smaller institutions, in comparison, have to be adept at navigating a number of these varied skill-sets. In either case, the process is expensive and time-consuming, and therein constrained by time and money.

Colati ended his talk with a number of questions about the nature of archives and special collections, and whether it is possible to rethink some of the assumptions that come with current archival systems. These included: Where should we invest our resources to improve access to and preservation of primary sources? Should we change where

---

[9] Jae Jennifer Rossman, "Investigating the Perceived Value of Special Collections in the Academic Library," *Journal of Library Administration* 60, no. 6 (2020): 631-644.

archives and special collections departments are located in the library system? How do the economics of the archives and the library system differ?

General Discussion

The general discussion for this session touched on a number of different topics related to the funding of Archives and Special Collections as well as their place within the library system. The participants were placed into three groups at random, and were all given the same questions to prompt discussion.

One of the major topics of discussion was that of revenue generation through digital reproduction of collections. Before COVID-19, some institutions did charge for digital reproductions as a means to create revenue, although it was often used as a means to fund the digitization specialists. Some institutions did not pursue collection digitization even before the pandemic, while others chose not to charge for digital reproductions. There were various reasons for this, including institutions with large endowments not needing revenue generation, worries that setting up infrastructure for payment would outweigh the potential earnings, and philosophical reasons such as not charging students, faculty, and notable alumni. Some institutions had also charged for reproduction in the past to slow the amount of requests they received. However, during the COVID-19 pandemic, it was almost ubiquitous across institutions to not charge for digitization requests.

Another major topic of discussion was the process of digitization from a financial and administrative perspective. The process of digitization is labor intensive and requires both planning and upkeep. More traditionally, archives and special collections have taken a "one touch" approach to their collections, meaning that an item should only be handled once in the digitization workflow. This means that it would be digitized at high quality and fully catalogued with detailed metadata. Some institutions are beginning to consider other options for the digitization process.

Lower quality images are often taken by students, faculty and archivists to serve as reference material. Some of the institutions present were interested in aggregating these "reference quality" images into some sort of repository that could then be searchable. The reference request would form the metadata associated with the image. This sort of process would add little to the current workflows of archivists and digitization specialists, but would make their collections more accessible. However, there was some trepidation that managing another repository, even with small-scale upkeep, would add difficulty to current practices. One proposed solution is to create inter-institutional networks to share the burden of maintaining the infrastructure. To achieve this, institutions would need to adopt a more universal approach to their collections, in that they would create a shared network of information. The Big Ten Academic Alliance fairly recently took a step in this direction with their agreement in regards to the borrowing and lending of archives and

special collections materials through InterLibrary Loan (ILL). However, creating a network of reference quality images goes beyond ILL. Placing images into a single repository that would be searchable to all with access to any of the affiliated institutions would require further development in the ways that institutions approach issues like copyright and access to their collections. In this way, archives and special collections could better approximate an Open Access approach to their collections.

D. Scholars

The third topical session, "Scholars and Researchers," was held November 6, 2020. Organized by Wes Hamrick, Postdoctoral Fellow at Greenhouse Studios, the session brought together three researchers, at different career stages, including a PhD student in art history, an early-career historian, and a more senior historian with over two decades of experience conducting archival research. Panelists included Grace Kuipers, PhD student in Art History at the University of California, Berkeley; Hilary Bogert-Winkler, Assistant Professor of Liturgy at the School of Theology at the University of the South; and Mills Kelly, Professor of History and Director of the Roy Rosenzweig Center for History and New Media at George Mason University.

This panel discussion was designed to give workshop participants greater insight into the way that scholars have been conducting archival research since the start of the pandemic, especially as many archives remain closed or have highly restricted access, and as restrictions on travel necessitate conducting archival research remotely. The session was also intended to give participants a better understanding of how scholars conduct archival research more generally. For example, a common assumption among archivists is that scholars need box-level access to collections, while Hamrick's discussions with scholars reveal that researchers often have item-level knowledge of collections gleaned from citations and notes in scholarly publications. Addressing this discrepancy is particularly important for our work in developing Sourcery, which in its current iteration is expressly designed to provide item-level access to collections through remote provision of scans. Replicating in-person archival research at the box level, however, presents many more challenges that remain to be solved.

Hamrick's investigation of scholars' research methods provides important insights for archivists and information professionals. Duff and Johnson's 2002 paper on the information-seeking behavior of historians in archives provides a detailed description of how ten historians conduct their archival research, which includes searching for specific known items.[10] At nearly twenty years old, it is unclear how well the paper's findings reflect current research practices, but the panel discussion confirmed that known-item research remains an essential part of how historians and other scholars use the archives.

---

[10] Wendy M. Duff, and Catherine A. Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives." *The Library Quarterly: Information, Community, Policy* 72, no. 4 (2002): 472-96. Accessed May 9, 2021. http://www.jstor.org/stable/40039793.

The session also suggested that archivists and researchers often have somewhat different understandings of the relationship between archival research and image reproduction. For instance, many archivists tend to view image reproduction primarily through the lens of digitization. While archivists of course understand that one-off image reproduction requests from scholars entail somewhat different quality control standards than those required for large-scale digitization efforts, they nonetheless tend to view reference scan requests as a second-best option for researchers. Unlike in formal digitization workflows, camera angles and lighting are not carefully controlled, metadata may be less accurate or incomplete, and the images provided are divorced from the context of the surrounding pages or documents in the box or folder. Nonetheless, lower-quality reference scans are often perfectly adequate to the researcher's needs. Indeed, each of our panelists cited examples of this particular use case during the workshop session. Rather than image quality and complete metadata, the primary need of scholars and researchers is fast, easy access to the images of documents they wish to examine. Quick, ready access to reference-quality scans is precisely the use case for which Sourcery is designed.

Given that many scholars tend to prioritize access over image and metadata quality, another theme that emerged from our panel discussion is the potential for making reference scans provided to individual researchers available online for other scholars who might need them later on. As with the discussion about high-quality digitization vs. one-off reference-quality scans, the discussion here also centered around questions about image quality, metadata and discoverability. Scholars are almost universally supportive of making reference scans available online, and while archivist opinion is mixed, the practice is gaining support. One benefit of the panel discussion, however, was bringing this possibility to light and to surface some of the questions that doing so might raise. Subsequent to the workshops, we learned that the Folger Shakespeare Library, one of our partner institutions for testing and developing the enterprise version of Sourcery, has already piloted such a project. In fact, among other things, the Folger will be exploring how to incorporate Sourcery-generated reference scans into its workflow for making images available online.

The panel discussion raised a number of issues that suggest that archival research as a whole is still undergoing a long period of transition from the pre-digital era to the age of smartphones and digital access. For example, funding for archival research is often structured to support week- or month-long trips to a single archive to conduct exploratory research at the box level. While this mode of research is still a mainstay of how some scholars use the archive, many researchers express a need for remote, item-level access to documents cited in other work or in a finding aid. The scholars on the panel also mentioned their frustration with restrictions imposed by the archive on sharing copies of research scans. They noted that sharing scans is a longstanding practice among researchers and that, in their view, archive policies on sharing images are ultimately

unenforceable. Indeed, panelists also discussed the increasingly common phenomenon of using social media to find and share digital images of archival material. This practice reflects not only researchers' willingness to share images, but also the fact that researchers often prefer quick remote access to a particular document over a research trip to the archive. This disconnect between archive policies on sharing and reproduction, and the way scholars use and share images from their archival research, raises important questions about the relationship between "policy" and "copyright."

In addition, the panel discussion confirmed that while scholars' demand for remote access to collections is virtually limitless, limits on archives' time, labor and funding act as constraints on how that demand can be met. As one archivist in the session pointed out, even where collections can be digitized, additional considerations must be given to server space and long-term digital curation. Thus, while improvements to infrastructure can save time and labor with respect to scanning and digitization workflows, new infrastructure can entail additional time and labor in their own right.

E. Closing Session

The Closing Session took place on November 30, 2020. The keynote speaker was Clifford Lynch, Executive Director at the Coalition for Networked Information. After Lynch's remarks, the group broke into three breakout rooms for an idea creation activity, then returned to the main room for a general discussion and question and answer period.

Summary of Cliff Lynch's remarks

Lynch discussed how the pandemic has forced institutions to think about resilience in supporting instruction and research. While the pandemic has demanded the whole globe plan for resilience, many organizations have been dealing with natural disasters for a while. For example, in the last few years, California has experienced rolling blackouts, earthquakes, and fires.[11] We will continue to experience challenges due to climate change. For example, travel will probably become more difficult and more expensive. In addition, many archives are grappling with issues of equity and opening up their resources to a wider range of users. Climate change and the need for equity will not go away once the pandemic is over. Money would help solve some of these problems, but our institutions are under financial pressure. We're seeing a gradual move away from library budgets dominated by purchasing expensive materials to making library and institution-owned materials more accessible.[12] However, inside-out collections still take a financial commitment.

---

[11] Melanie Mason, "Natural Disaster is Inevitable in California. And It Can Define a Governor's Legacy," *Los Angeles Times*, September 30, 2018.

[12] Lorcan Dempsey, "Library Collections in the Life of the User: Two Directions." *LIBER Quarterly* 26, no. 4 (2016): 338–359.

Lynch explained that, as a profession, we need to reorient the way we think about archives and special collections. We might start thinking about remote access first, with in-person access as a nice-to-have. This has implications for descriptive practices. Archives have developed creative ideas that mix local proxies with online access, an example being Zoom consultations. Online access raises questions about archival gate-keeping and the risks of putting possibly sensitive information online. Another change is that, with so many facilities shut down or closed to the public, student labor is not available. The labor of on-demand digitization, which might have been delegated to paraprofessionals before, is now being done by archivists and curators. We need to think about the right mix of staff.

Another issue raised throughout the course of these workshops is the different needs of individual researchers, individual Sourcerers, the university, the library, the archive, and the public. Through different methods, Sourcery peer-to-peer and enterprise are both trying to balance the needs of different actors. As a community, we need to try to understand the right balance, which will vary from institution to institution.

There are also systems issues. When items are opportunistically digitized, we might want to connect them to records in a collection management system like ArchivesSpace, but most discovery systems don't accommodate this well. And how do individual actors give their photographs back to the archives? While these digital surrogates might not be up to the highest standards, they are better than nothing. In addition, the archivists also get the information that a researcher was interested in that item.

Finally, Lynch mentioned that pilot institutions for Sourcery are going to extend internationally. As we come out of the pandemic, there will continue to be restrictions on international travel. Lynch acknowledges that he doesn't fully understand the international barriers to sharing digitized archival materials but thinks that the Sourcery project presents a good opportunity to explore and learn about those obstacles.

Summary of breakout sessions

Participants divided into three groups for the breakout sessions. Each group discussed a theme that arose during the course of the previous sessions.

The first group considered the question, "How can we create sustainable practices from emergency access practices? What are some practices we can build into a new version of 'normal'?" Takeaways from the group included:

- Institutions want to continue offering remote consultations, such as remote video reference and live chat.

- They also want to continue new forms of co-curricular outreach, like videos of collections and recorded talks. One archive found that twice as many people attended remote events as previous in-person ones.

- Participants discussed building sets of digital materials that can be used frequently in various courses.
- Discussion included ways to store and maintain reference reproductions for reuse, like an aggregated reference catalog that could even be part of larger combined collections like the Digital Public Library of America (DPLA)

- There was also discussion of staffing. During the pandemic, staff members from outside of archives have been brought in to help with transcription, metadata, and other remediation work. Can that continue after a return to campus? There was also mention of the limited sustainability of previous staffing levels in general.

The second group was asked to describe the ideal workflows for providing digitized material to researchers. What equipment and standards would be used? How would it fit into metadata and repository ingest workflows? How would you balance this work with other priorities? This group reported several points of conversation around these issues, including:

- A self-serve option in the reading room that would allow researchers to capture images of acceptable quality. There would need to be systems for ingesting those images in a digital repository, possibly involving close collaboration between reference, digitization, processing, and cataloging staff.

- Participants mentioned that more labor and funding is needed. We could refocus budgets from outside-in collections to inside-out collections.

- It can be difficult to balance the digitization of researcher requests with long-term digitization priorities, which might surface collections that document groups that have been historically underrepresented

The third group examined one of the recurring themes during these sessions, that researchers need remote access to scans of documents more frequently and in higher volume than archives are able to provide. Clearly this is a problem rooted in limited funding and available labor, but how do we solve it? The group reported out several points, including:

- One idea involved digitization collaborations and image sharing across institutions.

- There needs to be a national-level commitment to digitization rather than

  funding specific institutions or projects. In addition, we could consider digital infrastructure as a public good. There should be discussion about getting funding from federal programs and adjacent fields that have deeper pockets.

- There needs to be greater clarity on what level of digitization is required for access and what is required for publication and long-term preservation. Could there be stratification of digitization levels?
- There was discussion of bringing in more labor through paid internships and service learning in English and history, or cooperative programs between archives.

- The question was asked, have we under-thought the technology of scanning? Are there better options than cell phones?

III. Future directions for Sourcery

In January of 2021 we began discussions with our first partner institution, the Folger Shakespeare Library, about testing and developing an enterprise model of Sourcery for use by archives and collecting institutions. The beta of the enterprise version was completed in February and we have begun initial customizations of the platform for use by the Folger. Over the spring and summer of 2021, we will continue working with the Folger, as well as other institutional partners, including Northeastern University Library, UConn Archives and Special Collections, the Hartford Public Library, and possibly Harvard's Houghton Library. We will work with our partner institutions to integrate Sourcery into their existing image request workflows in order to determine what additional features and modifications are needed. A key part of this integration process will be to document the various image reproduction workflows at our partner institutions, as well as at other institutions. Not only will documenting the scan request workflow help identify inefficiencies and choke points, it will also help us better understand what time and labor costs are incurred with each scan request, including costs that have potentially not yet been accounted for. During this period of testing and iterative development we will work toward the development of a more robust version of Sourcery that can meet the needs of a broad range of collecting institutions.

In addition to developing a more robust image request workflow, we will also work to integrate the enterprise version of Sourcery with Tropy and Omeka. From the workshops we learned that "off-the-books" sharing of images is already a feature of the archival research landscape, and that both researchers and collecting institutions recognize the role that digital infrastructure can play in facilitating image sharing at scale. A number of institutions, including the Folger, are beginning to make reference scans requested by

individual researchers available online in a searchable content management system for other researchers. Omeka, which like Sourcery and Tropy is a platform of the Corporation for Digital Scholarship, is a content management system (CMS) expressly designed for displaying collections of images. Integrating Sourcery with Omeka would allow an institution to create an automated pipeline between image requests placed through Sourcery and the institution's CMS for placing reference scans online. Tropy, meanwhile, is a tool that individual researchers use to store and manage their personal collection of images for archival research. Though Tropy was designed to store images locally on the researcher's hard drive, Tropy developers are currently working to implement cloud storage for images. Integration of Tropy with Sourcery and Omeka would allow researchers to share images taken in the institution's reading room with the institution's CMS for publishing reference scans online. While these reference scans would not replace conventional digitization efforts, integrating Sourcery with Tropy and Omeka in this way could nonetheless greatly expand remote access to archival collections.

When the pandemic necessitated travel restrictions and the closure of archives in early 2020, we were forced to put the peer-to-peer model of Sourcery on hold while we pivoted to working on the enterprise model. We are taking a wait-and-see approach to the peer-to-peer model, with an eye to restarting in the fall of 2021, conditions permitting. We believe at this point that the peer-to-peer model and the enterprise model can work in tandem. From the archives professionals who attended the workshops, we learned that while well-funded institutions with robust imaging services may adopt tools like Sourcery as part of their in-house infrastructure, many smaller institutions with fewer staff resources cannot provide regular imaging services for remote researchers. For these institutions, the peer-to-peer model of Sourcery may provide a means of remote access to their collections. As the workshop discussions revealed, the success of both the enterprise version of Sourcery, as well as the peer-to-peer model, depends on figuring out where Sourcery facilitates efficiencies versus where it might put additional pressure on institutional resources. Even where Sourcery presents challenges to current institutional arrangements, we look to explore both how archival research might be reconfigured and how the challenges Sourcery presents might enhance the case for greater financial support for archives and special collections.

Bibliography

Big Ten Academic. "Big Ten Academic Alliance Principles and Protocols for Interlibrary Loan of Special Collections Materials." April 2018. https://www.btaa.org/library/reciprocal-borrowing/special-collections.

"Browse LUNA: Folger Digital Image Collection." https://luna.folger.edu/luna/servlet/FOLGERCM1~6~6

Conway, Paul. "Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas". The Library Quarterly Volume 80, Number 1 January 2010.

CNI. "What Happens to the Continuity and Future of the Research Enterprise: Looking to 2020-2021 and Beyond." 2020. https://www.cni.org/go/research-continuity-sept-2020-update.

Dempsey, Lorcan. "Library Collections in the Life of the User: Two Directions." LIBER Quarterly 26, no. 4 (2016): 338–359. doi: http://doi.org/10.18352/lq.10170.

Dimunation, Mark. "Everything We See Hides Another: Coping With Hidden Collection in the 21st Century Library." In Something's Gotta Give: Charleston Conference Proceedings, 2011, edited by Bernhardt Beth R., Hinds Leah H., and Strauch Katina P., 21-29. Purdue University Press, 2012. Accessed April 29, 2021. doi:10.2307/j.ctt6wq4sf.7.

Duff, Wendy M., and Catherine A. Johnson. "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives." The Library Quarterly: Information, Community, Policy 72, no. 4 (2002): 472-96. Accessed May 9, 2021. http://www.jstor.org/stable/40039793.

Mason, Melanie. "Natural Disaster is Inevitable in California. And It Can Define a Governor's Legacy." Los Angeles Times. September 30, 2018. https://www.latimes.com/projects/la-pol-ca-next-california-natural-disasters/.

Milligan, Ian. "Becoming a Desk(top) Profession: Digital Photography and the Changing Landscape of Archival Research." Google slides, AHA 2020. https://docs.google.com/presentation/d/1zLOTxijUgQhPy_xxKtuWBTpwdoDoRxoxNyXvMyHAFt8/edit#slide=id.p.

OCLC. "REALM project research." February 11, 2021. https://www.oclc.org/realm/research.html.

Rossman, Jae Jennifer. "Investigating the Perceived Value of Special Collections in the Academic Library." Journal of Library Administration 60, no. 6 (2020): 631-644, doi:10.1080/01930826.2020.1748437.

"Tropy: Research Photo Management." https://tropy.org/.

Appendix A: Number of Participants

Attendees Per Session
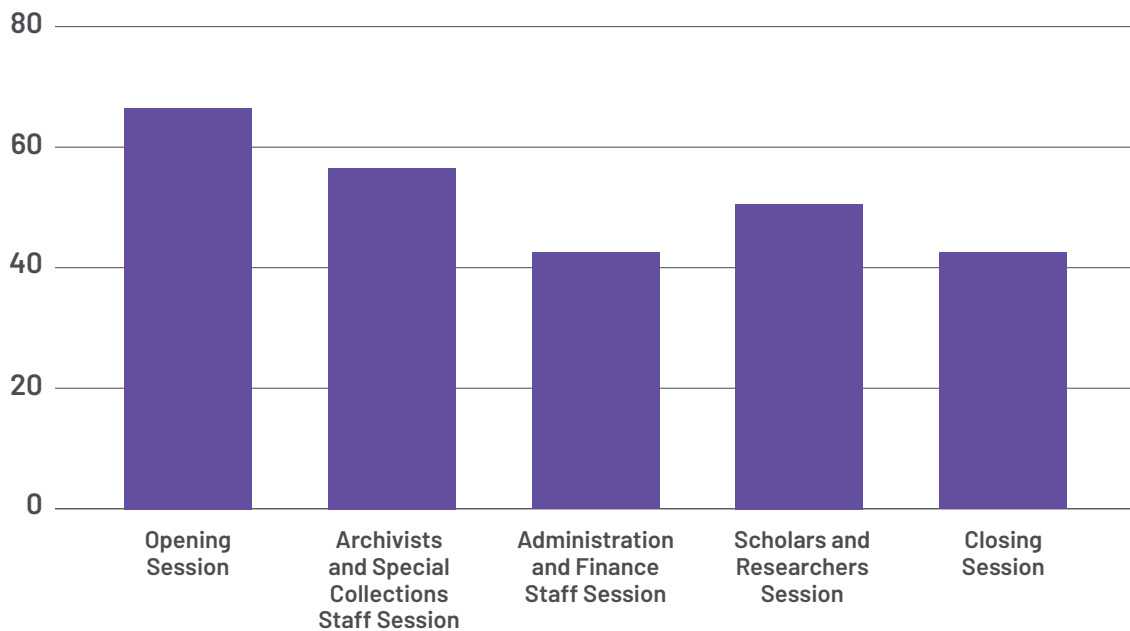
Opening Session: 67 participants

Archivists and Special Collections Staff Session: 57 participants

Administration and Finance Staff Session: 43 participants

Scholars and Researchers Session: 51 participants

Closing Session: 43 participants

Appendix B: Institutional Affiliations

American Antiquarian Society:
1. Kimberly Tony, Head of Reader's Services
2. Elizabeth Pope, Curator of Books and Digitized Collections

Amherst College:
1. Mike Kelly, Head, Archives and Special Collections
2. Tim Pinault, Digitization Coordinator
3. Este Pope, Head, Digital Programs

Andrew W. Mellon Foundation:
1. Patricia Hswe, Program Officer for Public Knowledge

Brown University:
1. Andrew Majcher, Head, Digital Services and Records Management
2. Amanda Strauss, Associate University Librarian

Coalition for Networked Information:
1. Cliff Lynch, Director

Dartmouth College:
1. Daniel Chamberlain, Associate Librarian for Digital Strategies
2. Jennifer Taxman, Associate Librarian for Research and Learning
3. Scout Noffke, Reference and Administrative Specialist

Harvard University:
1. Martha Whitehead, Vice President for the Harvard Library
2. Tom Hyry, Librarian of Houghton Library
3. Robin McElheny, Associate University Archivist for Collections and Public Services
4. Timothy Driscoll, Senior Reference Archivist
5. Dorothy Berry, Digital Collections Program Manager
6. Kate Donovan, Associate Librarian for Public Services

Jisc (United Kingdom):
1. Peter Findlay, Digital Portfolio Manager

Johns Hopkins University:

1. Winston Tabb, Dean, University Libraries
2. Liz Mengel, Associate Director Collections & Academic Services
3. Margaret Burri, Assistant Director for Academic Liaison and Special Collections
4. Alistair Morrison, Manager of Library Applications
5. Amy Kimball, Head of Materials Management
6. James Stimpert, Senior Reference Archivist
7. Jordon Steele, Hodson Curator of the University Archives
8. Katie Gucer, Manager of Digital Collections as Data
9. Ken Flower, Associate Director, Administrative & Discovery Services
10. Earle Havens, Nancy H. Hall Curator of Rare Books and Manuscripts
11. Gabrielle Dean, William Kurrelmeyer Curator of Rare Books and Manuscripts; Librarian
12. Sharon Achinstein, Sir William Osler Professor of English
13. Rebecca Brown, Professor, History of Art

Library of Congress:

1. Edith Sandler, Manuscript Reference Librarian
2. Chad Conrady, Archivist

Massachusetts Historical Society:

1. Elaine M Heavey, Director
2. Anna J. Clutterbuck-Cook, Reference Librarian
3. Kanisorn Wongsrichanalai, Director of Research
4. Daniel Hinchen, Reference Librarian
5. Hannah Elder, Reproductions Coordinator

Mount Holyoke College:

1. Leslie Fields, Head of Archives and Special Collections
2. Micha Broadnax, Digital Projects Archivist
3. Deborah Richards, College Archivist

Northeastern University:

1. Giordana Mecagni, Head of Special Collections and University Archivist
2. Molly Brown, Reference and Outreach Archivist

Princeton University:

1. William Noel, AUL for Special Collections
2. Jon Stroop, Director of Library IT and Digital Services
3. Alexis Antracoli, Assistant University Librarian for Special Collections, Technical Services
4. Sara Logue, Assistant University Librarian for Special Collections Public Services

Stanford University
1. Tom Cramer, Associate University Librarian
2. Benjamin Albritton, Rare Books Curator
3. G. Salim Mohammed, Head and Curator, David Rumsey Map Center
4. Mark Matienzo, Assistant Director for Digital Strategy and Access
5. Benjamin Stone, Curator for American and British History; Associate Director, Department of Special Collections

State Library of Massachusetts:
1. Dava Davainis, Head of Reference and Information Services
2. Judith Carlstrom, Technical Services Manager
3. Beth Carroll-Horrocks, Head of Special Collections
4. Silvia Mejia, Special Collections Librarian
5. Elizabeth Roscio, Preservation Librarian

Tufts University:
1. Alicia Morris, Assistant Director
2. Steve McDonald, Digital Initiatives Librarian

University of Connecticut:
1. Michael Kemezis, Repository Manager
2. Rebecca Parmer, Head of Archives & Special Collections

University of Massachusetts, Boston:
1. Jessica Branco Colati, Archives Track Program Director and Lecturer, History Department

University of New Hampshire:
1. Elizabeth Slomba, University Archivist
2. Emeline Dehn-Reynolds, Library Manager, Milne Special Collections & Archives

Wesleyan University:
1. Andrew White, University Librarian
2. Amanda Nelson, University Archivist
3. Suzy Taraba, Director of Special Collections & Archives

Yale University:
1. Jonathan Manton, Music Librarian for Digital and Access Services
2. Richard Boursy, Research Archivist
3. Melissa Grafe, Head of the Medical Historical Library
4. Barbara Rockenbach, University Librarian

Appendix C: Session Materials

Session Prompts:

I. Archivists and Special Collections Staff:

1. Tell us about how you are supporting distance researchers now:

    A. How have you traditionally provided services to remote patrons?

    B. What is the greatest challenge in providing access to collections now?

    C. What would make your life easier?

    D. Is your reading room open to researchers?

2. How are most of your collections described (finding aids, MARC records, inventories, accession records) and how does this affect researchers making requests?

3. Are there material types with particular challenges, such as AV material or born-digital collections?

4. Do you ever lend special collections through ILL? What is that experience like?

5. How do copyright concerns affect what you can and cannot digitize?

6. Do you have concerns about how remote access to collections is or is not providing more equitable access to archival resources?

7. What else hasn't been covered in these questions?

II. Administration and Finance:

1. How could/should the financial model be different? How could we imagine other ways to invest money in running Archives and Special Collections?

    A. What are positives and negatives in the current financial and process model of A&SC? For example:

        1. Expectation of revenue generation

        2. Expectation of fundraising/grant writing/ and donor stewardship to fund core activities

        3. Mediated access

    B. What alternatives could we develop? For example:

        1. New approaches to access

        2. Alternative funding streams from library budgets

        3. What new approaches to processing might we adopt—digitize first, describe later? Machine learning? AI metadata creation? Minimal description?

2. Has the current crisis changed your financial model?
Have these changes forced changes in how you do things?

    A. How do we approach access to collections moving forward?

        1. Are emergency activities, like free scans for example, a sustainable model?

        2. How do we account for the work that comes with increased digitization?

            a. Shift current staff to more digitization duties? What do we lose?

            b. Hire new staff?

    B. How do you see the financial model changing in the next year? 2 years? 5 years?

III. Scholars and Researchers:

1. Conducting Research During COVID-19: Broad Issues and Particular Challenges

    A. How has COVID-19 changed the way you do research (methods, timelines)?

    B. What are some particular challenges you've encountered or anticipate?

    C. Are you putting research on hold, exploring new lines of inquiry, etc.?

    D. What are some ongoing archival research challenges that pre-date COVID-19?

    E. Can you speak about the economics of archival research and equality of access/resources for researchers at different institutions or at different levels of seniority? In other words, to what degree is cost a factor in how you conduct archival research?

2. Remote Access after COVID-19: Opportunities and Solutions

    A. Is it possible to go back to "business as usual"? Recognizing that funding is always an issue, for both researchers and archival institutions, what things would you like to see done differently? (Researchers and archivists may have different perspectives here.)

    B. What kinds of services would help you conduct archival research remotely?

    C. For example, would remote, folder-level access be helpful? For the archivists, what would remote, folder-level access look like?

    D. If remote access to archives were more readily available, would you consider using more diverse/out-of-the-way collections in your research, beyond what is currently available?

Appendix D: Evaluation

After the workshop series, we sent out a survey to all registrants. Ten people responded. On a scale of 1–5, the average rating of the effectiveness of the series was 3.5. The session rated most effective was the closing session.

One participant commented that the workshop was "one of the most worthwhile things I've done this fall." Another respondent commented on the welcoming virtual space and well-organized program. We also heard that the speakers were "knowledgeable and capable." However, several participants commented about a lack of clarity around the content of the different sessions in terms of how Sourcery might facilitate access.

A majority of respondents would be interested in future sessions on the topic of remote access to archives and special collections. Overall, the participants were interested in Sourcery, but raised concerns about how it might integrate with existing archival workflows and systems and how it would impact the current researcher for hire ecosystem.

Another area of growth for the future is the diversity of the speakers and participants. One survey respondent commented that all the panelists were white, so that's something we would want to address in future sessions. In addition, many participants were from comparatively well-funded and prestigious institutions. Sourcery might be more helpful to those at smaller or less funded archives or colleges, so future sessions might seek participants from those institutions.