UNIVERSITY LIBRARIES

Library Faculty Publications

Library Faculty/Staff Scholarship & Research

7-26-2021

The Archives at the Tip of Their Fingers: Exploring User Reactions to Large-Scale Digitization

Emily Lapworth University of Nevada, Las Vegas, emily.lapworth@unlv.edu

Su Kim Chung University of Nevada, Las Vegas, sukim.chung@unlv.edu

Follow this and additional works at: https://digitalscholarship.unlv.edu/lib_articles

Part of the Archival Science Commons

Repository Citation

Lapworth, E., Chung, S. (2021). The Archives at the Tip of Their Fingers: Exploring User Reactions to Large-Scale Digitization. *Journal of Archival Organization* 1-55. Routledge. https://digitalscholarship.unlv.edu/lib_articles/721

This Article is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Article in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself.

This Article has been accepted for inclusion in Library Faculty Publications by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

The Archives at the Tip of Their Fingers: Exploring User Reactions to Large-Scale Digitization

Emily Lapworth^{a*} and Su Kim Chung^b

^a Digital Collections, University of Nevada, Las Vegas, United States; ^b Special Collections and Archives Public Services, University of Nevada, Las Vegas, United States

^{a*}emily.lapworth@unlv.edu, https://orcid.org/0000-0003-4412-2870; ^b https://orcid.org/0000-0002-7805-1911

Emily Lapworth is Digital Special Collections and Archives Librarian at UNLV University Libraries. Her work focuses on increasing online access to archival materials that document the history of Southern Nevada and gaming. She coordinates essential aspects of digitization, metadata creation and management, digital asset management, and digital preservation in order to provide ongoing public online access to thousands of unique historical items and to create the rich description that makes them easily discoverable and usable. Her current research agenda centers on better understanding user needs to further improve access to and use of digital archival collections. Emily earned her MLIS from Simmons University.

Su Kim Chung has been immersed in the history of Las Vegas since she began work in the UNLV Libraries' Special Collections Division in 1999. As head of public services, she is in charge of reference, outreach, and instruction for archival collections and photographs about Las Vegas and southern Nevada history. She is also responsible for conducting assessment about the in person use of Special Collections & Archives. Su Kim regularly provides instruction for students in how to conduct archival research, and has designed and led faculty retreats to encourage teaching faculty to incorporate research with archival collections into their assignments. She earned her MLIS and PhD from UCLA.

The Archives at the Tip of Their Fingers: Exploring User Reactions to Large-Scale Digitization

Advances in digital image capture technology and the adoption of More Product, Less Process methods have resulted in special collections and archives largescale digitization that creates a new kind of digital surrogate. Mirroring and reusing aggregate archival arrangement and description, these digital surrogates represent multiple items and are minimally described as a whole. The authors conducted interviews to explore user reactions to this digitization method at the University of Nevada, Las Vegas. This study found that large-scale digitization does have a positive impact for users, although additional strategies may be required to maximize the usefulness of the resulting digital objects.

Keywords: archives; special collections; digital collections; large-scale digitization; digital surrogates; users

Introduction

For researchers interested in using archival materials housed in remote repositories, digitization opens up new and exciting possibilities. Yet early digitization efforts typically only provided access to scattered documents in a collection, and without the ability to have virtual access to an entire collection, researchers were unlikely to view digital collections as a substitute for an actual research visit. The idea that researchers could have access to entire archival collections online seemed unlikely.

Digitization of library and archival materials was once a long and tedious process involving the use of flatbed scanners which a user had to open and close when scanning each individual document or photo. Not surprisingly, such devices limited the amount of material that could be scanned and put online. The possibility of digitizing entire archival collections seemed untenable due to the amount of time and staff it would take to complete this work. In the past ten years, however, advances in technology and equipment have made the large-scale digitization of archival materials and hence entire collections - a real and exciting possibility for researchers.

The introduction of large-scale digitization revolutionizes access to archival collections for users. It has the potential to enable virtual research that mimics the experience a user might have using a physical collection in an actual reading room. It also allows for new ways of doing research, such as computational analysis of large data sets created from digital surrogates. But how do researchers respond to viewing and using entire archival collections online? How does the presentation and description of digitized materials affect the way in which users engage with the content?

University of Nevada, Las Vegas (UNLV) Libraries Digital Collections, a department of UNLV Special Collections and Archives (SCA), began large-scale digitization efforts in 2015, upon acquiring a Digital Transitions Phase One cultural heritage rapid capture system. The Phase One system radically changed how staff were able to digitize SCA materials. With the assistance of grant funding, Digital Collections was able to apply large-scale digitization methods to four different archival collections. These collections were published online via the digital asset management system (DAMS), CONTENTdm. For the first time since UNLV Libraries began its digitization program, users can now access four manuscript collections online almost in their entirety. (Some formats were not digitized or were not digitized in their entirety, such as three-dimensional artifacts, newspaper clippings, magazines, and oversized materials.)

To explore how users reacted to UNLV's large-scale digitization efforts, the Head of SCA Public Services and the Digital Collections Librarian conducted semistructured interviews with researchers who have used UNLV SCA materials. These interviews provide insight into how users conduct archival research in both in-person and online environments, and how users interact with materials that have been digitized using large-scale methods and delivered via the CONTENTdm DAMS. This article will discuss the interviewees' reactions to UNLV Digital Collections' implementation of large-scale digitization and share recommendations based on that feedback.

Literature review

Large-scale digitization is characterized by the digitization of large quantities of materials using accelerated and efficient workflows. Rapid capture camera systems have increased the pace of digitization for cultural heritage institutions, creating high quality digital images in a fraction of the time it would take a flatbed or overhead scanner to do the same. (This article will discuss digitization that revolves around digital image capture rather than the digitization of audiovisual materials.) Optical character recognition (OCR) software can be used to quickly generate transcriptions from images of typewritten text, allowing for full text search.

Early literature about large-scale digitization pertained to the digitization of newspapers, magazines, and books.¹ Articles about the large-scale digitization of special collections and archives materials followed, mainly consisting of case studies or recommendations for how to successfully and efficiently implement it.² Large-scale

¹ Astrid Verheusen, "Mass Digitisation by Libraries: Issues Concerning Organisation, Quality and Efficiency," *Liber Quarterly: The Journal of European Research Libraries* 18, no. 1 (2008): 28-38.

² Examples include: Ricky Erway, "Rapid capture: Faster throughput in digitization of special collections," OCLC Research, 2011. Erik Moore, "Strategies for Implementing a Mass Digitization Program," *Practical Technology for Archives* no. 3 (November, 2014), https://practicaltechnologyforarchives.org/issue3_moore/. Craig Harkema and Cheryl Avery, "Milne En Masse: A Case Study in Digitizing Large Image Collections," *New Review of Academic Librarianship* 21, no. 2 (2015), https://doi.org/10.1080/13614533.2015.1034806. John Yolkowski and Krista Jamieson,

digitization depends in part on technology that allows for rapid capture, but it also requires a strategy for the efficient arrangement and description of digital objects for access and discovery, and it is this aspect that has the most effect on users.

The process for creating digital object metadata was traditionally item-level and intensive, but in 2010 Mark A. Greene suggested that MPLP (More Product, Less Process) should be applied to digitization, and digitized materials should be described at the file or even series level.³ The Digital Public Library of America (DPLA) Archival Description Working Group released guidelines for describing and representing aggregated objects in 2016.⁴ Arranging and describing digital surrogates in aggregate is more efficient than creating item-level metadata, but some institutions, such as the Archives of American Art and the Southern Historical Collection at the University of North Carolina at Chapel Hill, even went so far as to completely skip creating descriptive digital object metadata and instead provide access to digitized materials via archival finding aids.⁵

⁵ "Digitizing Entire Collections: Chapter 1, Background," Archives of American Art, Smithsonian Institution, accessed April 24, 2020,

[&]quot;Access and Preservation in Archival Mass Digitization Projects," *Practical Technology* for Archives no. 7 (December, 2016),

https://practicaltechnologyforarchives.org/issue7_yolkowski/. Cory Lampert, "Ramping Up," *Digital Library Perspectives* 34, no. 1 (Feb 12, 2018): 45-59,

https://www.emeraldinsight.com/doi/full/10.1108/DLP-06-2017-0020. Emily Lapworth, Sarah Jones, and Marina Georgieva, "Microfilm, Manuscripts, and Photographs: A Case Study Comparing Three Large-Scale Digitization Projects," *Journal of Contemporary Archival Studies* 6, no. 5 (2019), https://elischolar.library.yale.edu/jcas/vol6/iss1/5.

³ Mark Greene, "MPLP: it's not just for processing anymore," *The American Archivist* 73, no. 1 (2010): 194.

⁴ DPLA Archival Description Working Group, "Aggregating and Representing Collections in the Digital Public Library of America," (2016).

https://www.aaa.si.edu/documentation/digitizing-entire-collections-chapter-1-background.

Amanda Loeb conducted a content analysis of ten special collections and archives websites that have undertaken large-scale digitization and found that all of them provided some access to digitized materials via finding aids, and two of them exclusively used that method of discovery and access.⁶ Jane Zhang and Dayne Mauney also analyzed special collections and archives websites and they identified three different models for connecting digital objects to finding aids: 1) embedded, with the finding aid serving as the main interface for digital object discovery and access, 2) segregated, with multifaceted metadata serving as the main access point, but also providing an external link to the finding aid, and 3) parallel, which uses "both online finding aids and searchable item-level metadata to represent and display digital objects."⁷ Zhang and Mauney concluded that the parallel model offers the best of both worlds, and suggested that: "Archival description provides contextualized access to archival records, which may better serve research purposes. Item-level metadata provides granular access to information in the archives, which may better meet general information inquiry/use needs."⁸ However, the parallel model still consists of two

Tim West, Kirill Fesenko, and Laura Clark Brown, "Extending the Reach of Southern Sources: Proceeding to Large-Scale Digitization of Manuscript Collections," final grant report for the Andrew W. Mellon Foundation, University of North Carolina at Chapel Hill (2009).

⁶ Amanda Loeb, "Shopping the Online Archives Megastore: A Content Analysis of Special Collection Libraries and Archives Websites Produced Through Large-Scale Digitization," (Master's Paper for the M.S. in L.S. degree, University of North Carolina at Chapel Hill, 2013).

⁷ Jane Zhang and Dayne Mauney, "When Archival Description Meets Digital Object Metadata: A Typological Study of Digital Archival Representation," *The American Archivist* 76, no. 1 (Apr 1, 2013): 186.

⁸ Ibid., 190-191.

separate interfaces rather than the true integration of archival description and digital object metadata.

Multiple researchers have conducted studies to gauge user reactions to different models of digital object access (finding aid versus separate folder level and item level metadata). Tracy Jackson conducted a usability test of a parallel digital object access system and found that novice users preferred searching the item-level metadata, and advanced users reacted more positively to access via the finding aid.⁹ However, the existence of two separate interfaces was confusing for the novice users and some of the intermediate users.¹⁰ A 2012 usability study directly compared known-item searching by browsing a finding aid with links to digital objects, and by searching item level metadata.¹¹ Searching item level metadata took participants (mainly graduate and undergraduate students) less time and fewer interactions, but in this study the finding aid was not searchable, even using Control Find (ctrl+F). This limits the wider applicability of these results since many finding aids are (and it is recommended that they should be) keyword searchable. Joshua Ranger conducted a similar study comparing folder level and item level description in the same interface. When asked to complete 6 tasks, the "participants [graduate and undergraduate students] did not perform well with tasks related to the experimental [folder level] section, many failed or

⁹ Tracy M. Jackson, "I Want to See it: A Usability Study^[1] of Digital Content Integrated into^[1] Finding Aids," *Journal for the Society of North Carolina Archivists* 9, no. 2 (2012), 20-77.
10 Ibid., 68.

 ¹¹ Jody DeRidder, Amanda Presnell, and Kevin Walker, "Leveraging Encoded Archival Description for Access to Digital Content: A Cost and Usability Analysis," *The American Archivist* 75, no. 1 (2012): 143-170, https://doi.org/10.17723/aarc.75.1.5641v61p422u0u90.

gave up."¹² However, when informed that it costs much less to present the digital objects at the folder level, "ALL respondents stated that the Experimental Model WAS TO SOME DEGREE acceptable," e.g. "Better than not having it at all" and "Better than driving an hour away."¹³ DeRidder, Presnell, and Walker also noted the tradeoff of the cost effectiveness of the finding aid as interface approach. The Southern Historical Collection at the University of North Carolina at Chapel Hill conducted interviews and focus groups with their primary users (who have extensive experience with archival research) before undertaking the large-scale digitization of entire collections. Participants saw great value in digitizing entire collections and in accessing digital objects via the finding aids.¹⁴

Oya Y. Rieger recommended that archives and special collections privilege users in large-scale digitization, and consider context, structure, and materiality.¹⁵ Large-scale digitization of special collections and archives materials has been occurring for over 15 years, but there are many different methods and outcomes, which may or may not affect the user experience. Description, arrangement, presentation, discovery, and access are the major aspects of large-scale digitization that affect users and are summarized in this literature review. No two implementations of large-scale digitization are the same; however, the strategies and outcomes share similarities that allow for comparison across the archives and special collections landscape.

¹² Joshua Ranger, "More Bytes, Less Bite: Cutting Corners in Digitization," Society of American Archivists 72nd Annual Meeting, San Francisco, August 2008.

¹³ Ibid.

¹⁴ West, Fesenko, and Brown, "Extending the Reach," 33-34.

¹⁵ Oya Y. Rieger, "Enduring Access to Special Collections: Challenges and Opportunities for Large-Scale Digitization Initiatives," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 11, no. 1 (2010), doi://doi.org/10.5860/rbm.11.1.328.

Large-scale digitization at UNLV Libraries

UNLV Libraries Digital Collections department historically created richly-described online collections of digitized archival materials drawn from UNLV Libraries Special Collections and Archives (SCA), often with accompanying contextual websites. At the start of UNLV's digitization efforts there were no dedicated staff or strategies to create finding aids, and therefore no way to reuse metadata in an efficient way. The number of items within each digital collection ranged from about 200 to 2,000. The department was limited by the available technology they used, namely flatbed scanners and CONTENTdm digital collection management software. In 2015 UNLV Libraries acquired a Digital Transitions Phase One cultural heritage rapid capture system, which greatly increased the amount of items that staff were able to digitize in a project period. During this time a Technical Services department was also established in order to create finding aids for all archival collections. In order to keep pace with the newly increased digitization rate, Digital Collections changed their metadata strategy: instead of creating rich, item-level description, it focused on grouping multiple items together into compound digital objects, using existing arrangement and description from finding aids, and creating a minimal amount of original metadata for each digital object. (Compound objects may also be referred to as complex or aggregate digital objects.) Most of the metadata is created at the parent level to describe the entire grouping, and child items within compound objects are not described individually. (See Appendix C for screen captures of example single and compound objects from UNLV Digital Collections.)

UNLV's *Culinary Workers Union Local 226 Photographs* and *Entertainment* digital collections are the result of large-scale digitization methods. In both cases, the finding aid archival arrangement and description of the materials was mirrored and reused for the digitization process. In the *Culinary Workers Union* collection, each

physical folder corresponds to a compound digital object. Description from the finding aid was reused, and most of the digital objects do not have narrative descriptions, subject headings, or people or group names. In the Entertainment digital collection, digital objects correspond to the level of description in the finding aid. Most digital objects are equal to a physical folder of materials, although some costume drawings are described individually, as they are in the finding aid. Subject headings were added, and people and groups were identified using controlled vocabularies, but most of the metadata is at the parent level, and narrative descriptions of digital objects were not created. Computer-generated transcriptions were created using OCR software but were not corrected by staff, resulting in varying degrees of accuracy but still providing full text search capabilities. At the time of the user interviews, the Culinary Workers Union digitized materials were only available via CONTENTdm, but the digital object metadata contained a link to the archival collection finding aid (Zhang and Mauney's segregated model). The Entertainment digital objects were available via CONTENTdm (with a link to the archival finding aid in the digital object metadata) and also via links to the digital objects in the finding aid at the file level (Zhang and Mauney's parallel model).

Methodology

In 2019, the Head of UNLV Special Collections and Archives (SCA) Public Services and the Digital Collections Librarian conducted semi-structured interviews with nine SCA users. These interviews were part of a broader research project to better understand UNLV SCA and Digital Collections users, and to generate recommendations for improvement.¹⁶ The implementation of MPLP processing and large-scale digitization were internal changes that resulted in users accessing collection materials in new ways. The need to enhance access to born-digital materials and a project to design and implement a new DAMS were additional changes that provided the opportunity to incorporate user feedback into SCA's growth. The interviews therefore addressed multiple aspects of research into UNLV SCA and Digital Collections, but this article will focus on the findings related to large-scale digitization.

Although not representative of UNLV Digital Collections' entire online audience, the interviewees were selected to represent four categories of user groups who typically conduct research in Special Collections (Table 1). These included undergraduate and graduate students, teaching faculty, and community users. Of the nine interviewees, all were local residents, except one who was a frequent out-of-town visitor to the SCA reading room. All of the interviewees were selected because they had conducted both in-person and virtual research with SCA materials, but they had differing levels of experience with archival research. Some had extensive experience using UNLV Digital Collections (see Appendix A for more information about interviewees' experience).

¹⁶ The UNLV Office of Research Integrity - Human Subjects determined the study to be exempt. Interviewees provided informed consent to participate.

Table 1. In	terviewees.	Unless	otherwise	indicated,	all interv	riewees	were l	ocal	and
affiliated w	vith UNLV.								

Interviewee	Description
Interviewee 1	Former graduate student (architecture), current community user
Interviewee 2	History professor (out-of-town)
Interviewee 3	Journalism and Media Studies professor
Interviewee 4	Museum curator (non-UNLV)
Interviewee 5	Graduate student (history)
Interviewee 6	Undergraduate (history)
Interviewee 7	Undergraduate (public health)
Interviewee 8	Community historian (non-UNLV)
Interviewee 9	Undergraduate (history)

The interviews consisted of 35 questions and addressed topics related to both physical and online research with archival materials in UNLV SCA and UNLV Digital Collections (see Appendix B for questions). All interviews were conducted in person and recorded with a digital recorder. The interviews were transcribed manually and using Otter.ai.

The first section of the interview gathered demographic info, and sought to gauge the respondent's (self-described) level of experience with physical research conducted at UNLV SCA as well as special collections and archives in general. Interviewees were then asked about the subject matter of their research, and to indicate (from a provided list) the type of material they consulted during their reading room visit. The interviewees were also asked to discuss their familiarity and understanding of finding aids as archival access tools, and to select (from a provided list) the ways they used the archival materials during their research. The final question in this section queried users as to how the research experience in UNLV SCA could be improved.

The questions in Section II were designed to mirror those in Section I as a way to understand the similarities and differences users felt in conducting physical and digital research with archival materials. Users were also asked to view six different digital objects in UNLV Digital Collections: "Does the digital object and its description make sense to you? Tell us what you're looking at. Please share your thoughts by saying them all out loud [think-aloud protocol]." These pre-selected objects included items digitized and described at the individual level and objects created using largescale digitization methods: richly described single objects from the Photograph Collections, a richly described compound object from the Southern Nevada Jewish Heritage Project, and minimally described compound objects from the Culinary Workers Union and Entertainment collections. (Screen captures of some of the digital objects are included in Appendix C.) Responses provided insight into how users interacted with the system, their research process, their understanding of different levels of object description, and how they utilized digitized materials in their research. Although the responses from interviewees provided insight into how large-scale digitization can positively transform archival research, they also made it clear that the presentation of digital objects can significantly affect how users find, understand, and use digitized archival materials.

The authors read through and coded the interview transcripts multiple times. During the initial read-through, the authors took notes separately to identify themes. Next, answers to straightforward questions were compiled into a spreadsheet to allow for easier comparison across interviews. Additional analysis of the interview transcripts was also completed to identify and document in a spreadsheet mentions of specific system functionality, metadata, and finding aid elements. Another read-through focused on identifying themes relevant to large-scale digitization, including user reactions to compound objects and minimal metadata. Direct quotes related to large-scale digitization were compiled in a separate document and sorted into themes.

Although the small size of the interview group and the specific focus on UNLV prevents this study from drawing broad generalizations, it is a starting point for further research. These interviews contribute to a better understanding of how users respond to the online presentation of digitized materials, and how large-scale digitization methods could be improved.

Findings

The interviews revealed some common themes that illustrate how these users engage in the research process with digital collections, and the challenges they face when working with a digital asset management system that provides access to large quantities of digitized materials via compound objects. Their responses also provide insight into how research strategies used with physical archival collections transfer into the digital environment. The themes are described below, along with sample direct quotes from interviewees.

Entirety of collection available online

Prior to the implementation of large-scale digitization methods, researchers only had access to small digital collections of select materials curated by librarians. Digitizing entire collections provides access to more materials, and potentially more diverse materials. Interview responses suggested that the amount of material online has the potential to transform research projects. With entire archival collections online, researchers feel they have the opportunity for more in-depth research of their own choosing. Prior to the digitization of entire collections, most people were just using Digital Collections to find photos, browse thematic digital collections, or to view a

sample of what is available in the reading room; now online research can actually replace in-person research. Community users who may not usually come to the reading room also benefit from increased online access.

- "I always kind of see it as... a springboard for the original collections right? I can do my, more like the preliminary, but I guess now it'd be more close to the entirety [of my research] if I'm able to get, you know, whatever copies [online]." (Interview 5, question 26)
- "So it looks like about 99% of the items here have been digitized that can be viewed online. One can complete an extraordinarily large research project from this... There's no reason for there not to be many articles and parts of books or entire monographs using this collection, this is extraordinary." (Interview 2, question 22.6)
- "I really like the digitized pamphlets because...having this access to something like this deep as opposed to just the photographs definitely augments my, you know, whatever I might be researching." (Interview 5, question 22.5)
- "I'm looking at, you know, these cultural things that don't come up anywhere else in the historical record. Okay, that's kind of interesting. No one talks about that." (Interview 5, question 28)
- "I just know that it would make a lot of the collections accessible to a lot of the community... A lot of them may be intimidated to come into the university or have never stepped foot on a university. So knowing that they have the archive at the tip of their fingers would be of great benefit to them if they want to do some historical research or learn more about Las Vegas and their city." (Interview 6, question 35)

Efficiency of research

Not surprisingly, one of the strongest themes to emerge from the interviews is the efficiency that large-scale digitization can bring to the research process. With more materials available online, less time and travel are needed for in-person consultation of materials, particularly for out-of-town researchers. Some of the interviewees noted that increased online access to collections also gave them a significant opportunity to strategize for future in-person research work.

- "For me personally, it doesn't affect me because I'm local, but I'm also lazy and the less I can do without having to go in somewhere, the more I can do in my pajamas, the better." (Interview 1, question 26)
- "It would be remarkably more efficient. I would do that [online research] all before I came to Special Collections. Research time would be dramatically reduced, cost of research would be reduced." (Interview 2, question 26)
- "It would make it possible to do, of course, everything at home that previously you'd have to be going around for. It does make it faster in a lot of ways because you can copy and paste and save items as opposed to laboriously taking notes of everything." (Interview 5, question 26)
- "But for non-local folks. I think that is a big deal. You know, for people that are doing this kind of research that live in Oregon or something, that's a big deal." (Interview 4, question 26)
- "It would be most helpful because I would be able to not have to make a long trip over here and stuff the parking meter with quarters." (Interview 8, question 26)

- "It would make it much easier. I wouldn't have to stop researching once Special Collections closes. I can continue researching at home." (Interview 6, question 26)
- "If I was in a hurry, or if I did not utilize my time properly, this would have been a great tool, especially because Special Collections has [scheduled operating] hours." (Interview 9, question 23.2.1)

Even for local researchers, who do have the ability to conduct in-person research, there is the feeling that 24/7 online access would have a transformative effect on the research process itself, perhaps moving it away from its linear structure. It would enable a web of interconnected research tasks not possible with the physical materials. All nine interviewees said they would download, save, or print items from UNLV Digital Collections. Downloading was also mentioned 14 other times during the interviews by eight of the interviewees, and printing was mentioned nine other times by five interviewees. Six interviewees mentioned requesting high resolution reproductions a total of ten times (CONTENTdm only allows direct downloads of lower quality access copies). Seven interviewees said they would share, publish, reuse, or edit digital materials, and two said they would use Digital Collections for a digital humanities project (such as computational analysis of text, visualization, etc.). One interviewee also discussed his desire for UNLV Digital Collections to be compatible with Zotero, to allow easy citation management.

• "And I could explore, I can download. I could print. I could read. I can move around, you know, internally, in the computer, go to other collections, get leads." (Interview 8, question 26)

- "I have downloaded a lot and saved and printed and it's great when I could just you know, as a graduate student I'm working rapidly on something, to be able just to get an idea about what's out there and download the links or images that I can refer back to later." (Interview 5, question 28)
- "I would know that everything that I could access that's accessible is available so I could do my research. I could explore options. I could give examples to other people to look at, we could split the labor if I'm working with a couple of graduate students. I could access from home or outside the country without fear of having 'Oh, I've got to really follow up and do a last check to make sure that I didn't miss anything that was in the actual physical collection.'" (Interview 3, question 26)
- "But I think the bottom line is if it was all, if every single thing was available for me to view online, it would be much easier to constantly reference the collection for whatever I was working on." (Interview 4, question 26)

Online access offers advantages not available via in-person research, such as the ability to instantly download, print, and share items. If OCR software is used to create transcripts of textual materials, then full text searching can also make the research process easier and more efficient. Full text search of transcriptions was mentioned 16 times by seven of the interviewees.

• "Yes, I do [search the full text of digitized documents]. I have done that. I'll give you an example from the African American collection. I was doing research on the March 26th, 1960 integration agreement, and there were so many people who were critical to that, I would just type in names and see if they would show

up on any documents in that collection. That was very, very helpful." (Interview 2, question 31)

• "The transcripts were very, very useful to find stuff." (Interview 5, question 22.6)

Collection-based research strategy

The research behavior of the interviewees revealed that most of them relied on a collection-based research strategy when working with digitized materials. They sought out related materials by looking for other items in the same archival collection, as opposed to items with the same assigned subject terms. Seven out of nine interviewees clicked on the archival collection in the digital object metadata, with many of them mentioning multiple times how the archival collection was helpful in their research to find similar items. Archival collection metadata was mentioned 33 times total by interviewees when they viewed the sample digital objects and discussed their thoughts. This is likely due to the fact that almost all of the interviewees had experience doing inperson archival research, which is collection-based. Other digital collections users who are not familiar with the concept of archival collections would probably have different search and browse strategies.

- "These ones like 'women' are really unhelpful to me. 'Families,' unhelpful. Most all of these, these are rarely helpful, the item subject, when they're that general, like 'Families.' Graphic elements are generally unhelpful." (Interview 1, question 22.3)
- "I will branch out once I've found something that's in the ballpark by clicking on the archival collection, or metadata like category that's specific to that building or location... Those tend to be the two most frequent metadata

categories that I will use as I'm going through Digital Collections, the collection itself, and something specific to that item." (Interview 1, question 22.2)

- "This is helpful too, because if I'm looking for more images then it has the link to the original collection, then sometimes, I'd find stuff that I wouldn't even know I was looking for, like, 'Oh wow...look at this collection, there's all sorts of other things that are relevant to my interest."" (Interview 5, question 22.1)
- "It tells you the original collection, which is again really important for, maybe you could find, you could click on it, which is great because you could go directly to the Freeman Collection, and maybe this collection has more photographs or letters or newspapers that could potentially add to the research." (Interview 9, question 22.1)

A collection-based research strategy is compatible with large-scale digitization strategies such as digitizing entire folders and collections, including links to digital objects in finding aids, and including links to finding aids in digital object metadata. Mirroring physical archival research digitally appears to be helpful to researchers who are already used to it, but it may not be helpful to those who lack the same experience.

Finding aids

The finding aid is a staple access tool that researchers use when doing archival research in the reading room. It is not surprising that interviewees familiar with finding aids from conducting in-person research also found them helpful in the digital environment. Those interviewees who were not accustomed to using finding aids still found them essential when they viewed compound objects that lacked the traditional rich description of individually digitized objects. Interviewees felt that the finding aid provided important contextual information for minimally described digital objects. All nine interviewees agreed that a link from the digital object metadata to the finding aid was very helpful, and it was mentioned a total of 24 times by interviewees. Five interviewees mentioned that the "front matter" of the finding aid (collection level description and notes) in general is helpful. Specific elements of the finding aid that were mentioned as helpful include: the biographical/historical note (mentioned by six interviewees), the scope and contents note (mentioned by three), the arrangement note (mentioned by four), and the inventory (mentioned by seven).

- "I would probably click on this [the link to the finding aid] to find out more about the collection it's involved with." (Interview 3, question 22.1)
- "Wow, this is great. Everything the researcher wants, you know who's in the image, you know the place, you know the date, you know how to put it in context of the other documents by just clicking on that [the link to the finding aid]." (Interview 2, question 22)
- "[The finding aid biography] really gives you an idea of why, you know, why it is that we're documenting and talking about and saving the material of this guy." (Interview 4, question 14)
- "The background information on those [finding aids] are very, very useful too, and just those little synopses, that gives you an introduction to the topic...so I'd say those are essential for sure." (Interview 5, question 14)
- "Yeah, so now that I know about the [finding] aid, I can just go to the historical note and read it [the contextual information] there." (Interview 6, question 23)
- "This is exactly what I was like when I mentioned contextual evidence, having a resource like this is exactly what I meant. Like, there's an abstract that details what a large majority of the photos are, historical context." (Interview 7, question 22.4)

 "Definitely, the finding aids are very, very useful. In just getting an idea even instead of going straight to searching the image databases." (Interview 5, question 14)

Placing a digital object in the context of the finding aid and the archival collection was seen as valuable to interviewees. Two interviewees wanted the link from the digital object to the finding aid to go to the exact spot in the finding aid where that digital object "is;" currently, these links just bring users to the first page of the finding aid. Another idea mentioned was to visualize where something is in a collection, similar to how seeing a physical collection in person gives one an idea of where an item is in a collection, as well as the collection's extent, formats, etc. Eight interviewees expressed a desire to know the extent of either the compound object or the entire collection.

- "So this isn't telling me where this [digital] object is within the finding aid... what is likely is seeing that image and wanting to know what else is around that image. And so I would come to the finding aid not to look for that image but to see what other images are also related to it." (Interview 1, question 23.1)
- "Maybe I missed it but I didn't see how that didn't take me back to where it was listed in the collection guide... Cause it would be nice if you could be, 'oh, I love that image, I wanna go back and see some other things right after that."
 (Interview 2, question 22.6)
- "Going up one level, I'm trying to get a sense of the landscape of this collection, and so having that initial view which you can really get in a physical environment, you can get a sense of you know, how many documents are here, what types of documents do I have, what's the overall impression of what's

going on. Is there any way to display that, when you go into a collection? Just conceptually." (Interview 3, question 22.3)

Although large-scale digitization ostensibly allows for entire collections to be digitized, interviewees were still concerned when searching through these collections online that they could be missing something housed in the physical collection. Notifications added into the digital objects that indicated when folders were not digitized in their entirety were thus appreciated by researchers.

- "I have a tendency to think that I missed something online. And if I'm, if I have the file folders in front of me, I know that I've gone through everything."
 (Interview 4, question 32)
- "What I'd be most worried about is any omissions, so something that might be poorly scanned or improperly scanned, it's like with Google books, every once in a while I see a page that's not scanned properly, or something might be left out." (Interview 3, question 33)
- "I've never seen this notification before that the item has not been scanned in its entirety. That's really helpful to know." (Interview 1, question 22.5)
- "I like your idea of making a statement, in that, one of those first lines [of the finding aid], that says how much of this particular archive is available online or has been digitized, so that you're warned right up front that these are the things that I won't have access to online, just so I know that I wasn't missing something. Because that was the fear." (Interview 4, question 33)

In that respect, finding aids with direct links to digital objects at the folder level were seen as a tool to help ensure that researchers aren't missing something (Figure 1). The interview responses also revealed a general appreciation for the folder-level one-toone links as a tool that made the research process easier and more efficient: eight interviewees mentioned how useful these were a total of 17 times. The preference of experienced archival researchers for access to digitized archival materials via the finding aid was also evident in the feedback gathered by the Southern Historical Collection at the University of North Carolina at Chapel Hill for their large-scale digitization project in 2007-2009, and in Tracy Jackson's 2012 study.¹⁷ The Literature Review subgroup of the Digital Library Federation Assessment Interest Group User Studies Working Group also identified "clear benefits when digital library content is either directly embedded into archival finding aids or linked to from the finding aid. Additionally, the more granular the descriptive content in the finding aid is, the more the digital content is used."¹⁸

Title/Description	Containers		
Awards and honors: Dance Educators of America, press clippings and telegram, 1967 May (view online)	box 1	folder 1	
Awards and honors: Greater Las Vegas Chamber of Commerce EPY (Entertainment Personality of the Year) award, press clippings and programs, 1976 January (view online)	box 1	folder 2	
Biographical information: magazine articles and press clippings, 1950s-1980s	box 1	folder 3	
Biographical information: press releases, 1960s-1970s (view online)	box 1	folder 4	

Figure 1. Screenshot of UNLV SCA finding aid with folder-level embedded links to digital objects.

• "The viewing online [folder-level links to digital objects in finding aid] is the thing, that's super cool. That would be super helpful. Especially for me who

¹⁷ West, Fesenko, and Brown, "Extending the Reach, 33-34. Jackson, "I Want To See It, 72.

¹⁸ Joyce Chapman, Elizabeth Joan Kelly, Liz Woolcott, and Tao Zhang, "Surveying the Landscape: Use and Usability Assessment of Digital Libraries," Digital Library Federation Assessment Interest Group User Studies Working Group, (2015), doi:10.17605/OSF.IO/9NBQG.

says, 'ah, I want to see what's around this image.' Then you go into the finding aid, see the stuff that's around it, and then see the stuff that's available to view online." (Interview 1, question 23.2.1)

- "I really like the links [in the finding aid]. This would make research so much easier." (Interview 6, question 23.2.2)
- "That [having the links to digital objects in the finding aid] makes the research process so much quicker." (Interview 8, question 22.4)
- "This [finding aid] is a great tool. This whole document is just amazing. It goes directly to it [the digital object]. Very helpful." (Interview 9, question 23.2.2)
- "Collection guide. This is what I used to find exactly what I needed when researching Donn Arden....This is so important. And I love that it's connected directly to the digital [objects]." (Interview 9, question 23.2.1)

Physical versus digital

Although interviewees frequently emphasized the convenience of using digitized objects in the research process, they still retained a strong affinity for the physical materials. In some cases, it was for a practical reason: a need to view an original costume design to see the attached fabric swatches, or the importance of viewing the physical condition of an original photograph as an art historian or exhibit curator. The fear of missing something came up again, this time on the item level. Small details such as captions written on the back of photographs were important to researchers, indicating the importance of reproducing items as fully as possible, or at least documenting any omissions in the metadata of the digital surrogate.

- "The other thing that I wasn't able to tell from the online files was if the Jackson sketches were originals or copies, and so that was another reason why it was important to come in to see it in person." (Interview 4, question 28)
- "I know that there are some times where there are things written on the back of a photograph, and you can see the indent of words, you know, on the back, and there have been times where I've come in to look at the backside of a photograph to see if stuff's written on it... if the digital image says 'these words were on the back' [if the words on the back of a photo are transcribed in the metadata], cool, great. I trust that." (Interview 1, question 33)
- "For example, if there were photographs [with captions of the back] that were actually transcribed [in the metadata], I might want to see what is actually on the back. I did a paper a while ago on images that had been damaged through use over time, they hadn't been properly archived, and so I was really interested in the patina and the textures on it and that stuff is not really gonna come through [digitally], so yeah, I wanna be able to mark something for yeah, I wanna follow up [and view it in person]." (Interview 3, question 28)

In other cases, however, the desire to view the original physical materials represented an almost sentimental attachment to the physical object or the space itself. This is in line with the findings of Anastasia Varnalis-Weigle, who conducted a study comparing the user experience of interacting with physical objects and their digital surrogates.¹⁹

¹⁹ Anastasia S. Varnalis-Weigle, "A Comparative Study of User Experience between Physical Objects and Their Digital Surrogates," *Journal of Contemporary Archival Studies* 3, no. 3 (2016), https://elischolar.library.yale.edu/jcas/vol3/iss1/3.

- "I do enjoy seeing the physical objects. The tactile aspects, if I can actually touch them, how they smell, just all these other senses are really important I think." (Interview 3, question 33)
- "I think I always would prefer just to sit with a collection physically because there's something really fantastic about touching things and flipping through things, the experience is really satisfying." (Interview 4, question 32)
- "I do think that some items would be really fun to see in person. I guess I'm a history nerd." (Interview 6, question 33)
- "Sometimes it's just nicer to see something in person. You know, like for the Union Pacific Collection, to see the actual documents and letters, you know, actually hold them in your hand. It's a different feeling." (Interview 8, question 32)

Still, it was interesting to note that most of the interviewees did not perceive any fundamental differences between the online and in-person research process. As previously discussed, the differences that were mentioned mainly revolved around how online access is easier and more convenient: it allows for quick references to materials and there are enhanced features such as full text search, downloading, sharing, etc.

- "I don't know, in my mind, they're all research whether it's digital or physical." (Interview 1, question 32)
- "I don't think [there's] much [difference], again I would just use the finding aid and do what I did when I was physically there opening up each folder by hand and turning these documents over." (Interview 2, question 26)
- "I think I use it [the digitized collection] the same way [as the physical]." (Interview 4, question 31)

- "I would use them [the digitized collections] the same way I use them in person." (Interview 6, question 31)
- "I wouldn't say [there's] any difference [between the physical and digital] that matters to research." (Interview 7, question 32)

Appreciation for physical SCA space and archival expertise

Although they clearly recognize the value of large-digitization in making more archival material available online, some interviewees still indicated an appreciation (and in some cases a preference) for the physical space of the reading room, as well as the expertise of the archives staff in providing guidance in the use of collections. As noted above, these feelings are related to the physical experience of working with collection materials and a near reverence for the reading room space itself.

- "I feel like it's more of an environment, like you're coming up here to do
 research...I like coming up here...but having it online....like if you're
 researching and when special collections isn't open, that'd also be really nice. I
 would like them both together." (Interview 7, Question 26)
- "The social aspect of learning might be affected detrimentally [by researching online], because I'm just figuring out everything on my own as opposed to talking with other people who have maybe spent their lives researching the topic." (Interview 5, question 26)
- "Part of it is having your expertise at my avail." (Interview 4, question 33)
- "It is different in the sense that in-person is usually more directed, directed research, and of course you find stuff in person too that you didn't know existed, and would want to maybe follow a new avenue of research on that, but I would

say that it seems to be more focused in person as opposed to online." (Interview 5, question 31)

• "I go more in depth when it's physical. So I could spend hours here in this space. It's an amazing space, you're surrounded by history. In this space, I could spend hours and I would look extensively at each folder, at each box that I found relevant to my topic." (Interview 9, question 31)

Trusting digital content

Although assessing the integrity and authenticity of digital materials has long been of concern to archivists, and an essential part of digital preservation training, interviewees expressed implicit trust in UNLV's digitized materials as accurate representations of the physical objects. This was largely based on their trust and confidence in the archival staff and their expertise. All nine interviewees said that they trust the digital representation of archival materials, although, as shown above, some discussed their desire or specific need to view the original materials in person.

- "I don't have any distrust at all of the...maybe I should, but I don't have any distrust, I think it has to do with having confidence in the archivists. If you've used an archives for many years, you just inherently trust the people who are there and you just assume that they're going to have an accurate representation in the digitized version." (Interview 2, question 33)
- "I wholly trust the digital collections. This is someone's profession. I have faith in that." (Interview 7, question 33)
- "I trust, my need to go to the physical collections has nothing to do with not trusting the digital database. It has everything to do with how I research and how I want to be immersed in my topic." (Interview 9, question 33)

 "I trust that the representations are done very well...I'd say we have a lot of trust in, I guess it's implicit trust in the institutions that they're not falsely representing something. So yeah, I definitely have a very, very high degree of confidence, or trust of everything that's put on the Special Collections website." (Interview 5, question 33)

Understanding CONTENTdm compound objects

Interviewees were shown several examples of compound objects in CONTENTdm that are equivalent to a physical folder of archival materials (Figure 2). In the center of the compound object page the current child image is displayed, taking up most of the screen. To the right, child items are displayed as small thumbnails. Below is the metadata.

When interviewees interacted with these compound objects, they did not always understand the structure and how it was supposed to reflect the physical arrangement of the original folder. It did not stop them from engaging with the material, but it did result in some confusion and misunderstanding. Those interviewees who had more experience with traditional archival research seemed to have an easier time understanding this concept, especially when accessing the compound objects from the finding aid folderlevel link. When interviewees were shown the compound objects in CONTENTdm first, without the context of the finding aid, there was not an initial recognition that the images were part of the same folder. Terms like "list," "collection," and "series" were used to refer to the compound objects, highlighting the challenge of labeling digital objects and archival components in the digital environment. Archival terminology is often unfamiliar, and familiar terms may have multiple meanings.



Figure 2. Screen capture of the first compound object shown to interviewees (question 22.3). The red box shows how much of the page is visible when first opened on a desktop computer. (Appendix C contains an enlarged version of this image, and other examples of digital objects shown to interviewees.)

- "Oh that's interesting, what's going on here?" Sometimes these throw me off when there's multiple images in a...on the side." (Interview 1, question 22.3)
- "Well this is a new feature, this is one in a series, and all the ones in the series are clearly identified on the right side. The first two [examples of digital objects] were single images... but I liked this over here. You're not having to keep clicking on the full collection to know what other photos exist of her or this family, that's good." (Interview 2, question 22.3)
- "Ok, this is sort of interesting here. Is there, ok, it wasn't immediately clear to me that this image was part of a larger list over here on this side." (Interview 3, question 22.3)
- "Was the other one like this? I didn't notice all the other places that you can, all the other objects and images...That's super helpful. I mean, the more on one screen I think the better." (Interview 4, question 22.3)

Compound objects in CONTENTdm feature one level of description on the parent object level, which is displayed above the description of the child item. Some interviewees appeared confused by the two levels of description. Some interviewees did not even scroll down the page far enough past the parent object level description to see the item level description (the screen capture above shows how much of the page was visible to interviewees when it was first opened).

• "Ok, so this short description is related to that, that's all the information we have on that. One of the reasons that I'm struggling a little bit with this sort of layout is I'm used to using something like Lightroom, which has the object, the visual object in the center, it's got metadata on the side, and then at the bottom it's got a list of all the objects, so this sort of standardized layout that I think a lot of people who work with visuals are familiar with." (Interview 3, question 22.6)

• "The object description and the item description is very confusing for the larger collections [compound objects]. So maybe renaming them collection description versus like, individual description [would help]." (Interview 9, question 35)

Interviewees suggested creating titles that clearly identify the digital object as consisting of multiple items, or as a folder, and listing the number of items within it.

- "See, this is what I'm a little bit more used to looking at with the multiple items because it will say something like 'folder 4 of 17' or 'images,' you know, 'negatives, 1 of 8' or whatever." (Interview 1, question 22.4)
- "This is a photo from folder 4 of 17 folders. That's a nice thing to know when you click it on. That you're not looking at a solitary piece of information. And then I like this over here, and it's clear here that there are gonna be many more pages." (Interview 2, question 22.4)
- "So I think it can again be a little confusing. However this is probably one of my favorites because it's so specific, it tells you what the photographs are, specifically what it is, where, when, what folder it's out of, how many images are actually in this collection [compound object]." (Interview 9, question 22.4)
- "Could you list how many items are in this folder [compound object]? I assumed it was more than one because it's [the title is] plural." (Interview 4, question 25)

Once the interviewees understood that compound objects were organized based on folders, most indicated a preference for searching this way. Six interviewees said that they preferred digital "folders" (compound objects) over individual items, one said he had no preference and would work with whatever is available, one wanted both, and one preferred items. Once interviewees were told that more items would be available online if they were organized into folders, all of them agreed that folders were acceptable. This is in line with Joshua Ranger's findings in a similar study.²⁰

- "I would prefer by the folder this way, for correspondence, for Las Vegas City Council meeting minutes. If it comes together, like physically, I'd prefer to see it as an object together physically." (Interview 1, question 24)
- "I would prefer to browse the folders...that would enable me to simply browse that particular folder. It would simulate the physical turning of the pages by my hand. That would be my preference." (Interview 2, question 24)
- "I would prefer to search on a file folder that said, for instance, Donn Arden Jubilee, and then see everything in that particular folder." (Interview 4, question 24)
- "I liked the folder because then it had related things. The other ones had like links to similar things like the courthouse and other pictures of courthouses, but this had like directly linked material which was really cool." (Interview 7, question 24)
- "[I prefer] folders, I would see the amount of materials that were in that folder and say, you know, this looks like all of these are good, or some of these might be good. If I saw just one individual, I would probably want to go to the folder anyway to see the others." (Interview 8, question 25)

²⁰ Ranger, "More Bytes, Less Bite."

 "The groups of things I think would be more helpful, especially in terms of like photographs, because you don't want to look through individual photographs." (Interview 9, question 24)

One interviewee (a journalism and media studies professor) noted that he would like to be able to view digital objects as both individual items and within the structure of folders, and he would like to organize the digital objects into his own folders:

- "It'd be nice to be able to have both [folders and individual items], I don't see why both wouldn't be possible. (Interview 3, question 24)
- "It's your folder, right, it's not my folder...I wouldn't be able to translate that
 into my folders of stuff. Maybe I'm interested in, not the specific people or
 whatever organizational scheme you used to create this folder, but I'm interested
 in outdoor architecture in Las Vegas in this era and I want to pull pictures from
 multiple collections and drop them together and create my own folders. But then
 I want to see how they link to your folders." (Interview 3, question 22.3)

Images as compound objects

The presentation of images as compound objects in the CONTENTdm digital asset management system provided some challenges to users during the searching and navigation process. On one hand, compound objects allow users to look through large numbers of related images, and they allow Digital Collections to provide online access to more images. On the other hand, compound objects also hide many images from search results in CONTENTdm. Individually described images presented as single objects can be browsed as thumbnails on the search results page. When images are within compound objects, users must navigate to the compound object in order to see all child images. Compound object children do not appear in search results, only the compound object parent record and a single thumbnail image do. Large-scale digitization makes textual records more accessible even if they are displayed as compound objects because OCR-generated transcripts allow for full text search, but the large-scale digitization of images does not result in the same additional benefits to users. As Grace Therrell demonstrated in her 2018 study, lesser levels of description hinder the discovery and retrieval of digital images.²¹ One interviewee, who is a self-described "power image looker" who uses UNLV Digital Collections mainly to find images, highlighted the problems of images as compound objects. Another interviewee suggested a possible compromise though: create a rich description for one representative or sample child item within the compound object, which would make the entire compound object more discoverable.

- "It would be really hard for me to find that [photo of a specific building] if I were looking for it because I had to get to the fifth [page of items within the compound object]...it's not in the metadata. And I understand that, I'm not... I've researched enough buildings through this to know that that's kind of how it is when researching background stuff, that you're not gonna get it, you're not always going to get it in your metadata." (Interview 1, question 22.4)
- "I would just think the challenge for me as a researcher then is how do I, when I get into this [compound] object and I look and I see there's 12 pages of, or 116 individual pictures to look at, how do I search within this to find that one page

 ²¹ Grace Therrell, "More Product, More Process: Metadata in Digital Image Collections," *Digital Library Perspectives* 35, no. 1 (2019): 2-14, https://doi.org/10.1108/DLP-06-2018-0018.

that I'm actually looking for. That would be my challenge as a researcher." (Interview, question 25)

- "I only look at individual photos and so I'd prefer to see each individual photo, individually, individually marked, that would help me most as a researcher"
 (Interview 1, question 24)
- "Is it possible to take essentially a folder of materials, and do something in detail about something that's sort of representative of what's in it, like choose a picture or sort of an iconic element from that, and then have the rest of the folder that's not as well described, so it could be sort of, I guess I'm trying to do a little bit of each." (Interview 3, question 25)

Another possible solution would be to allow users to add their own descriptive tags to digital objects and items. Four interviewees said they would add their own keywords or tags to digital items if possible, two said yes to private tags but no to public, one said "it depends," and two interviewees said that they wouldn't add tags.

- "Yeah. Because maybe I would understand it differently than the title described it... I could maybe explain that to someone else." (Interview 7, question 30)
- "I think it depends, I mean a lot of these areas I don't feel like I'm enough of an expert on these topics... I think that could easily be abused though... so I'd be a little worried. Crowdsourcing is really problematic." [Added that he would like to be able to add a citation if an item was used in a peer-reviewed academic article] (Interview 3, question 30)
- "I would be careful. And the reason why I say that is because, private, yes. Own tags, your own digital research, that is in your own private sphere. Sure. If the

public has access to the collections and is updating certain things... Publicly, No. Privately, Yes." (Interview 9, question 30)

Discussion

The researchers interviewed in this study were almost all experienced in conducting inperson archival research using finding aids, and they indicated a desire to replicate the same research process virtually. Having access to entire archival collections online greatly expands the amount and depth of research that users are able to accomplish virtually. Interviewees appreciated the advantages that online research has over inperson research, and they trust the digital representation of archival materials, although they did express an attachment to the physical experience, and in some cases, a concrete need to see an item in person when certain qualities could not be reproduced in the digital surrogate. Interviewees expressed a fear of missing something in their research, so clear information about what is digitized and what is not is important to them.

These interviewees who are experienced in in-person research continue to value a collection-based research strategy and finding aids. Even less experienced researchers said that the finding aids were very helpful for providing contextual information about the digital object, collection itself, and its creator(s). For digital objects with minimal metadata, a direct link to the finding aid was especially useful. Folder-level links in the finding aid to digital objects aid greatly in collection- and finding aid-based research.

Although interviewees were confused by the compound objects in CONTENTdm that were equivalent to an entire physical folder of materials, they were not discouraged. On the contrary, most interviewees said they preferred browsing entire "folders" of digitized objects, and all interviewees agreed that folders were acceptable if it meant that more materials would be available online. The compound objects' relationships to the physical arrangement was more obvious when interviewees navigated from a link in the finding aid to the digital object, rather than from the digital object to the finding aid.

In addition to the links between the finding aid and the digital object, the ability to search full text transcripts created with OCR software also mitigates the lack of descriptive metadata that often results from large-scale digitization. Images, on the other hand, do not benefit from this enhanced access. The ability to quickly browse large amounts of images online is a positive result of large-scale digitization, but when images are aggregated into compound objects in CONTENTdm, this hides them from search results and requires users to navigate to individual compound objects to view the child images. Combined with minimal descriptive metadata, this results in compound objects of images being less discoverable than single objects of images. However, compound objects of images are still more accessible and easier to browse then physical folders of images described at the same level. The same can be said of any format digitized using large-scale approaches that reuse archival arrangement and description; although these minimally described compound digital objects, they are more accessible than the original physical materials.

Based on the themes and feedback evident in these interviews, the authors compiled the following recommendations for the implementation of large-scale digitization approaches that mirror aggregate archival organization and description:

- Undertake usability testing to make your digital collections website as easy to use as possible. Implement user-friendly page layouts, labels, and visual cues.
- Provide resources such as online tutorials to help researchers use your website, and to help them understand your collections. Users accustomed to in-person

research may need help adapting to online research, and online users who have never conducted archival research in person may be unfamiliar with archival concepts and terms.

- Make the relationship between the original physical materials and the digital surrogates as explicit as possible. If finding aids are located separately from digital objects, provide direct persistent links between the two.
- Make it clear to users what has and has not been digitized. If a folder is digitized, document if it was fully or partially digitized, and which items were skipped.
- Represent items as fully as possible in the digital environment. For example, if there is anything written on the back of a photograph, transcribe it in the metadata, or digitize the back.
- Use OCR software to generate searchable text transcripts.
- Identify efficient ways to improve digital object metadata, especially for images.
 - Consider enhancing the descriptive metadata of one sample representative image within an aggregation, in order to make the entire aggregation more discoverable.
 - Consider allowing users to contribute tags to digital objects. Separate and clearly label the tags as user contributions.
 - Explore developing technologies for computer-generated description, such as image analysis and facial recognition.
- Support compatibility with existing research management and citation software.
- Explore possibilities for presenting digitized items both individually and within archival aggregations, in order to support use by researchers who are and aren't familiar with in-person archival research.

• Explore possibilities to support enhanced reuse of digitized materials, such as user-created collections and aggregations; computational analysis of data sets of digital surrogates, metadata, or text transcripts; and other digital research methods.

Conclusion

Our interview findings suggest that while large-scale digitization makes more content available online, the manner in which the material is presented in the digital environment can make things tricky for researchers. They may struggle with understanding how digital objects reflect the physical items or the organization of the physical collection, and how to use the content management system efficiently. However, they have trust in the digitized items and they see great potential in the opportunity for online research to replace in-person research of the same materials. The distinction between individually digitized items and items digitized via large-scale digitization methods is less important than the simple fact that more materials are available online. Description and context is important to researchers, and the finding aid is central in providing the kind of context that is lacking in minimally described digital objects. Some users would like to be able to move beyond the organization of digital surrogates that mirrors the archival organization of the physical collection (and the finding aid); they would like to organize the digital materials into their own research folders and collections. Despite the fact that large-scale digitization makes more material available online, users still enjoy the physical experience of being in the reading room working with the materials, and have an appreciation and respect for the expertise of the archivist.

This study examined an implementation of large-scale digitization that produced minimally described compound digital objects (also known as aggregate or complex digital objects) that repurpose and mirror the arrangement and description of archival finding aids. This study was limited to UNLV SCA users who had experience with inperson archival research and most of whom were comfortable using finding aids. Additional research should be conducted to compare the findings of these interviews with the attitudes of researchers using other digital collections, representing other implementations of large-scale digitization. Digital collections users who are not familiar with in-person archival research are also an important user group that deserve more attention. Minimally described aggregations of digitized archival materials can also be compared to minimally described aggregations of born-digital archival materials; as born-digital archival materials multiply, the digital collections access environments that were designed for individually described digital items become less useful for both archives and their users.

Appendix A: Interviewee profiles

Table 2. Interviewee profiles. Unless otherwise indicated, all interviewees were local and affiliated with UNLV.

		UNLV Special Collections and Archives (in-person research)		UNLV Digital Collections		
Interview number	Interviewee description	Self-described expertise: 5 is expert and 1 is novice	How often used?	Self-described expertise: 5 is expert and 1 is novice	How often used?	
1	Former graduate student (architecture), current community user	4	Several times per year	5	Several times per year recently; previously monthly or more	
2	History professor (out of town)	4	Several times per year since 2003	3	Almost daily in 2017 and 2018	
3	Journalism and Media Studies professor	3	2-3 times per year	3	2-3 times per year	
4	Museum curator (non UNLV)	1	Not regularly but intensively (1-2 times, 8 hours each, in the last few years)	2	Every other month	
5	Graduate student (history)	2	Once per year	4	Not regularly but intensively. 2 months for 3+ hours per day	
6	Undergraduate (history)	3.9-4	Monthly (at least once per month, 2-3 hours per visit)	2	Not regularly but intensively; maybe once per year	
7	Undergraduate (public health)	1.5 or 2	Once per year	3	Several times per year	

8	Community historian	3	Four times per	4	Constantly
		_		-	
9	Undergraduate (history)	5	Several times per month	3	Several times per year

Appendix B: Interview Questions

Collecting data for user-focused improvements to digital collections

Co-principal investigators: Emily Lapworth and Su Kim Chung, UNLV Libraries

The purpose of this study is to collect information about current and potential users of UNLV Digital Collections via an online survey and in-person interviews. The information will be used to inform improvements to UNLV Digital Collections. The results and methods will also be discussed in an article to be submitted to a peer-reviewed journal to be shared with the libraries, archives, and cultural heritage institutions community. The methods will be shared so that they can be reused by others, and the data and results will be shared to inform profession-wide understanding of digital collections users.

The data that will be used for this study will be anonymized. Interviews will be recorded (audio only) and transcribed. The audio will be deleted after the study period (2 years). All data will be stored in secure electronic folders on secure servers at the University Libraries.

This interview will have two parts. In part one we will ask you about your research and how you use UNLV Special Collections & Archives in person. In part two we will ask you to look at UNLV Digital Collections and answer related questions.

[Obtain informed consent from participant.]

[Ask participant to verify or correct the information collected from their reading room registration form]:

- 1. Organization or agency:
- 2. City, State:

- 3. Who are you? (Select all that apply)
 - College or university faculty or instructor
 - Department:
 - Graduate student
 - Department:
 - Undergraduate student
 - Major:
 - Teacher (K-12)
 - Subject:
 - Student (K-12)
 - Other:
- 4. Subject of research:
- 5. What is the output of your research? (Select all that apply)
 - Class paper or project
 - Teaching
 - Digital humanities project
 - Thesis or dissertation
 - Scholarly article
 - Newspaper, magazine, or online article
 - Book
 - Exhibit
 - \circ Film or video
 - Professional report or presentation
 - Social media, blog, or personal website
 - Genealogy research
 - Other:

Part I:

- 6. Please describe your experience conducting research in archives. (e.g. How many years? How often?)
- 7. What other archives do you conduct research at, besides UNLV Special Collections and Archives?
- 8. How often have you used UNLV Special Collections and Archives materials in person? (How long?)
 - Monthly
 - Several times per year
 - Once per year
 - Not regularly but intensively
 - Other:
- 9. On a scale of 1 to 5, how would you rate your expertise as a user of UNLV Special Collections and Archives? 5 would be expert and 1 would be novice.
- 10. Can you give us a brief summary of the research you've conducted in UNLV Special Collections and Archives?
- 11. What specific collections or subjects have you researched at UNLV Special Collections and Archives?
- 12. What kinds of special collections materials do you generally use/have you used? (Select all that apply)
 - Photographs
 - Oral history transcripts
 - Oral history audio or video
 - Audiovisual materials (not oral histories)
 - Manuscript collections
 - Maps
 - Born-digital records (materials that were created digitally, rather than physical materials that were scanned or digitized)
 - Other
- 13. Beyond viewing materials, what else do you do with them? (Select all that apply)
 - \circ $\,$ Take pictures, scan, or photocopy to refer back to later

- Request high-resolution reproductions
- Share, publish, or reuse
- Cite
- Other
- 14. When doing research in UNLV Special Collections and Archives, how do you use the finding aid, collection guide, or other descriptions? (front matter, etc.)
- 15. Is there anything specific that makes research at UNLV Special Collections and Archives challenging? Any improvements you would like to see or suggestions you have?

Part II:

- 16. How confident do you feel conducting research using library and archives databases? (for example online catalog, quicksearch on libraries homepage, Special Collections database)
- 17. Have you viewed archival materials online before? (where, what, etc.)
- 18. Have you used UNLV Digital Collections?
- If yes, the participant has used UNLV Digital Collections:
 - 19. What did you use UNLV Digital Collections for?
 - 20. How often have you used UNLV Digital Collections?
 - Monthly
 - Several times per year
 - Once per year
 - Not regularly but intensively
 - Other:
 - 21. On a scale of 1 to 5, how would you rate your expertise as a user of UNLV Digital Collections? 5 would be expert and 1 would be novice.

All participants:

22. Go to the following webpages and look at the UNLV Digital Collections materials. Does the digital object and its description make sense to you? Tell us

what you're looking at. Please share your thoughts by saying them all out loud [think-aloud protocol]:

o 22.1

http://d.library.unlv.edu/digital/collection/pho/search/search/search/pho017 826 [digital object from Photograph Collections (1 of 2)]

o 22.2

http://d.library.unlv.edu/digital/collection/p17304coll4/search/searchterm /pho025998 [digital object from Photograph Collections (2 of 2)]

o 22.3

http://d.library.unlv.edu/digital/collection/jhp/search/searchterm/jhp0004 52 [digital object from Southern Nevada Jewish Heritage Project]

• 22.4

http://d.library.unlv.edu/digital/collection/cwu/search/searchterm/cwu02 43 [digital object from the Culinary Workers Union Local 226 Las Vegas, Nevada Photographs]

o 22.5

http://d.library.unlv.edu/digital/collection/ent/search/searchterm/ent0006 20 [digital object from Entertainment collection with 15 items, some items not fully digitized, and OCR transcription in metadata]

• 22.6

http://d.library.unlv.edu/digital/collection/ent/search/search/term/ent0008 88 [digital object from Entertainment collection with 116 items and OCR transcription in metadata]

- 23. Go to the following webpages and click on the link next to "Collection Guide." Does the collection guide change your understanding of the digital materials?
 - o 23.1

http://d.library.unlv.edu/digital/collection/cwu/search/searchterm/cwu02 43 [digital object from the Culinary Workers Union Local 226 Las Vegas, Nevada Photographs]

- 23.2 Digital objects from the Entertainment collection:
 - **23.2.1**

http://d.library.unlv.edu/digital/collection/ent/search/search/search/rem/e nt000620 [digital object from Entertainment collection with 15 items, some items not fully digitized, and OCR transcription in metadata]

 23.2.2 http://d.library.unlv.edu/digital/collection/ent/search/sea nt000888 [digital object from Entertainment collection with 116 items and OCR transcription in metadata]

- 24. Would you prefer to browse folders of materials online (similar to the reading room experience) or would you rather browse individual items?
- 25. Taking into account that more materials would be available online if they were grouped together in folders, does this change your answer?
- 26. If collections were available online in their entirety how would this affect your research?
- 27. What kinds of file formats would you like online access to?
 - Image files such as jpeg
 - High resolution image files such as tiff
 - PDF documents (full text searchable)
 - Plain text documents (full text searchable)
 - Compressed (smaller size) MP3 audio files
 - Uncompressed (larger size but better quality) WAV audio files
 - Other:

28. With UNLV Digital Collections materials, would you:

- Download, save, or print so you can refer back later
- Share, publish, reuse, or edit
- Use for a digital humanities project (such as computational analysis of text, visualization, etc.)
- Cite
- Find materials that you later consult in person at UNLV Special Collections & Archives
- Just look
- Other:
- 29. Would you gather materials into your own virtual collection if possible?
- 30. Would you add your own keywords or tags to digital items if possible?
- 31. Would you use UNLV Digital Collections the same ways you use UNLV Special Collections and Archives materials in person, or differently?

- 32. Compare the experience of using a physical collection and using the digitized online version. Full collections available online include:
 - Culinary Workers Union Local 226 Photographs
 - Sands Hotel Public Relations Records
 - Donn Arden Papers
 - Jerry Jackson Papers
- 33. Do you trust the digital representation of archival materials? Or do you still feel the need to see the physical items for yourself?
- 34. Do you know how to cite digital collections materials? Would you cite the digital surrogates or the original physical materials?
- 35. What about UNLV's Digital Collections can be improved?

Appendix C: Screen captures of sample single and compound objects shown to interviewees



Figure C1. Digital object pho017826 from the Photograph Collections (1 of 2), shown as part of interview question 22.1. This single object is an item that is presented individually and richly described.



Figure C2. Digital object jhp000452 from the Southern Nevada Jewish Heritage Project, shown as part of interview question 22.3. This compound object contains multiple photographs described as a whole (at the parent object level), and also at the item-level; each child item has its own unique descriptive title.



Figure C3. Digital object cwu0243 from the Culinary Workers Union Local 226 Las Vegas, Nevada Photographs, shown as part of interview question 22.4. This compound object was created using large-scale digitization methods. It contains multiple photographs described as a whole (at the parent object level), reusing file-level description from the archival finding aid. Some of the parent-level metadata is repeated at the child item level. The child item titles are constructed from the parent object title, with the image (child item) number added.



Figure C4. Digital object ent000888 from the Entertainment collection, shown as part of interview question 22.6. This compound object was created using large-scale digitization methods. It contains multiple items described as a whole (at the parent object level), reusing file-level description from the archival finding aid. The only child-item-level metadata is the digital identifier. The item-level text transcript (created using OCR software) is also available on the page, but the section is collapsed in the screen capture.