Context Is (Almost) Everything: Academic Historians and Digital Archival Collections

Donald Force  
*University of Wisconsin - Milwaukee*

Bradley Wiles  
*University of Wisconsin - Milwaukee*

Follow this and additional works at: [https://digitalcommons.kennesaw.edu/provenance](https://digitalcommons.kennesaw.edu/provenance)

Part of the Archival Science Commons, and the History Commons

**Recommended Citation**  

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Provenance, Journal of the Society of Georgia Archivists by an authorized editor of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.
Introduction

Access to online archival materials has become vital for many academic historians and other researchers, especially during the COVID-19 pandemic. International travel restrictions, institutional lockdowns, and reduced service hours threaten the academic progress and research agendas of academic historians around the world. As a result, the pandemic has aimed the spotlight on the digital archival collections hosted by archives institutions and available through the web. However, even once the tide of the pandemic recedes, it is uncertain that academic historians will completely revert to their regular ways in accessing primary sources. Similarly, the fragile financial state of many archives may make it impossible for them to restore pre-pandemic services and staffing levels. In light of this increased role of digital archival collections, it is imperative for archivists to gain a better understanding of academic historians’ perceptions of these materials.

Previous research and analyses describe the complicated but symbiotic relationship between archivists and historians, a reality that continues to evolve as multidisciplinary collaborations around digital initiatives proliferate. Throughout the digital shift of the last three decades, digital archival collections have become invaluable to historians’ work, but original physical materials still “constitute the ideal” for research and teaching. The study reported in this article builds on previous work in the archives, history, and library and information studies disciplines to assess current perceptions that academic historians have toward digital archival collections.

We define digital archival collections as aggregations of digital archival objects, which may be born-digital records or reformatted electronic representations (i.e., digital surrogates) of archival materials that host institutions make available to users. Like their physical counterparts, a digital archival collection consists of mostly original and unpublished material resulting from everyday human and institutional activity. These virtual collections are similar to what are commonly understood as digital libraries, but like physical archival collections, digital archival collections are often more complex than digital libraries because of the unique relationships among the materials within each collection.

The survey findings indicate generally favorable perceptions of digital archival collections among the respondents, especially in terms of availability and access. This positive perceptual shift corresponds with respondents’ acknowledgement of higher rates of use and citation in recent years. This trend will likely accelerate in the wake of the COVID-19 pandemic and further digital development but increasing reliance on these materials also surfaces several issues around academic historians’ professional identity, research practices, and approaches to the historical record.

Research Questions

The survey instrument and subsequent analysis were based on two research questions: What perceptions do academic historians have of digital archival collections? How is this reflected in their use and citation of these materials? We intentionally made the distinction between use and citation. By use we mean that the respondent has referred to a digital archival collection or a digital object within such a collection and it assisted them with their research, but these were not cited in a publication. In other words, the digital archival collection or object served as a guidepost to other sources or served some other function rather than as direct evidence for the researcher. The survey asked two specific questions about use and the citation of digital archival collections and included a variety of questions concerning what perceptions academic historians have of these materials, especially when compared to original, physical objects, and what they believe to be the major benefits and limitations of the digital surrogates. The survey did not explore any one specific area in significant detail; instead, we sought to gain general insights about the perceptions that these users have of digital archival materials to identify areas for further investigation.

Literature Review

There is significant literature on the function, design, and content of digital libraries that informs our understanding of digital archival collections. While much of the general literature concerns the creation of and access to digital libraries or digital items from an institutional perspective, several recent studies on usability and information seeking behavior incorporate user perceptions. Similarly, archives scholars and practitioners have emphasized digital archival

---


collections’ use and utility through indirect means like citation and platform analysis, while also engaging directly with historians’ perspectives on their experiences with these resources.

In addition to the desirability of enhanced access and availability, several studies focusing on historians’ use of digital archival collections indicate the importance of context, serendipity, and archivist mediation in navigating collection interfaces and locating resources. These and other user studies indicate a clear preference among most academic historians to see and handle the original archival object. The hesitancy to use digital archival collections is often due to logistic and cognitive barriers (technological skills, collection knowledge, subject relevance, information overload, etc.), and an awareness that the lack of completeness and contextual information associated with online objects skews their interpretation of the materials. Studies by archivists and librarians point to user perceptions of reliability as a major driver of digital archival collections use or non-use—a quality that academic historians assess at the item, collection, system, and institutional levels.

Research and perspectives from the historian side generally align with what archivists have found about the reliability and function of digital archival collections. For example, in describing digitized newspaper databases, Richard Abel points to the inconsistency of text search results that seems to vary with each search attempt using the same keywords and phrases. Alexander Maxwell highlights the importance of presentation in digital archives, including tools

---


to manipulate and enhance digital objects and simple, intuitive interfaces that enable browsing; however, he mostly values digital collections as duplicative backups for the original (and more reliable) physical collections. Brian Ogilvie warns that “the interests of archivists, librarians and donors do not necessarily coincide with those of historians,” a reality with implications for a digital archival collection’s utility and academic historians’ perceptions of trust. Questions of trust seem to be at the heart of larger concerns that frequently underlies historians’ attitudes toward digital archival collections:

Despite worries of having too much material, historians are the ones most likely to want to digitize everything. Another rule of the profession is ‘save everything.’ One never knows when a piece of information has historic evidence. As much as historians might view the amount of material now available with a certain amount of trepidation, we also have emerged from a culture of scarcity that has preprogrammed us to not discard anything.

The inherent dichotomies associated with digital archival collections—scarcity and abundance, digital and physical, tradition and innovation, new and old—reflect a deep-rooted skepticism toward digital technologies within the academic historian profession. This likely has as much to do with the lack of historian involvement in system design as it does with the nature of historical research, which relies primarily on non-contemporary material that does not always translate well into the digital arena. A longitudinal study of historians across State University of New York campuses from 1992-1998 revealed the difficulty in new skills acquisition and problems with technological gatekeeping as the main impediments to historians adopting new digital practices that would ostensibly benefit their research efforts. The time and effort spent mastering novel technologies and methods necessarily diverts energy and attention away from those already tested and deemed reliable, especially regarding the source material on which historical scholarship is traditionally based.

Weller implicitly links the notion of trustworthiness to the various levels of context that historians must consider in approaching digital archival collections. Context is essential to determine a source’s provenance, to draw inferences from gaps and lacunae, to interpret its meaning, and to incorporate it into a convincing historical narrative. Weller insists that context is frequently overlooked or underemphasized in training new historians who increasingly take for granted disintegrated access to digital archival collections and other technological

---

15. Maxwell, “Digital Archives and History Research: Feedback from an End-User.”
enhancements that have tempered expectations in recent years. She contends that “a new generation of historical scholars are growing up without considering the difference the medium can make to the interpretation,” and suggests that knowledge of digital information provenance—a crucial type of contextual knowledge—should be cultivated along with other fundamental skills for research and teaching to enable historians to identify and select quality resources.

Lara Putnam also illustrates how the digital environment creates more opportunities for de-contextualization, which has implications for how historians approach their work when they lose the “multidimensional awareness” typically obtained through analog research. Advantages gained in the efficiency of resource retrieval, or in the erasure of time-cost barriers to access, must be weighed against evolving moral and methodological considerations for historians. As Putnam indicates, “Now you glance, you fish, you feast. But how much do you really know about the sources you find: about where they’re coming from, literally, politically, and culturally?”

The findings of the present study confirm that academic historians understand and acknowledge the importance of context in selecting and using digital archival collections. For this paper, we rely on the Society of American Archivist’s definition for context: “1. The organizational, functional, and operational circumstances surrounding materials’ creation, receipt, storage, or use, and its relationship to other materials. 2. The circumstances that a user may bring to a document that influences that user’s understanding of the document.” As we discuss in the following sections, our study respondents appeared to be less concerned with a digital archival object’s trustworthiness (i.e., its authenticity and reliability) than suggested in previous studies. Instead, academic historians have become acutely aware that the divorce of online digital archival collections and objects from the original context, materiality, and expert mediation has wide-ranging implications for their profession.

Methodology

We targeted academic historians at United States-based institutions designated as Doctoral Universities with very high research activity (R1) according to the Carnegie Classification of Institutions of Higher Education. We believed academic historians from such institutions would be likely to incorporate digital archival materials into their research process and have well-formed opinions about their value. We also believed that these historians would have significant experience in physical archives and would be willing to share their perspectives on differences between digital and physical archival collections.

According to the Carnegie Classification of Institutions of Higher Education directory, there are 131 R1 colleges and universities in the United States. A total of five institutions were excluded from the study: three do not have history departments, one did not explicitly provide

---

21 Ibid., 394.
24 Ibid.
email addresses for faculty members, and we excluded another because of faculty involvement in a related pilot study. From the remaining 126 institutions, we gathered the contact information for each faculty member by extracting their name, title or position rank, and email address into an Excel spreadsheet.

We only collected data from tenure-track and tenured faculty (e.g., Assistant Professor, Associate Professor, and Full or Distinguished Professors). We excluded adjuncts, teaching academic staff, emeritus faculty, and Ph.D. students because not all programs provided the contact information for these individuals. We included visiting faculty if their email was provided and they were either an Assistant, Associate, or Full Professor, regardless of their academic affiliation. In total, we compiled a list of 4,104 history faculty across 126 institutions.

We created a survey instrument in Qualtrics XM using questions and participant feedback from a pilot study we conducted in early 2020. The survey consisted of between 15 and 23 branch logic questions determined by the participant’s response. These contained mostly objective, closed-ended multiple choice questions, although several included an open-ended Other response option (see Appendix A). We distributed the survey via email on 11 August 2020 and sent a follow-up reminder to those who had not completed it two weeks later. We closed the survey on 6 September 2020. Of the 4,104 academic historians invited to complete the survey, 662 started it for a response rate of 16%. Of the 662 respondents who started the survey, 635 finished it for a completion rate of 96%, though not everyone who completed the survey answered every possible question.

Although the survey was directed at tenured and tenure-track faculty, we received responses from individuals with other personnel ranks and designations at the history departments we contacted. For the remainder of this article, the statistics provided reflect only the responses by Assistant, Associate, Full, and Distinguished Professors, which equates to 578 respondents, or 91% of the total (635). We combined the responses of the Distinguished Professor category into the Full Professor category on the basis that these ranks are functionally equivalent for our analytical purposes. The statistics are mostly descriptive and provide a quantitative basis around the respondents’ digital archival collection use and citation, which sets up the larger discussion on respondent perceptions.

Although over 600 academic historians completed our survey, our findings cannot be generalized to all academic or non-academic professional historians. In addition, our study cannot speak to the use of the materials by any other type of user, such genealogists, hobbyists, or teachers. Different types of users would likely have different perspectives on the value and utility of digital archival collections, all of which deserve consideration. The findings and observations presented in this paper provide only a glimpse of the data we collected from the survey, and we expect to follow other relevant routes of inquiry in future examinations.

Findings & Discussion

Participant Characteristics

25 When quoting the open-ended responses, we used an anonymous identifier (e.g., R1). We also made minor grammatical and punctuation edits to some of these responses for clarity.
The vast majority of the sample set (81%) were tenured faculty with associate or full professor rank. Assistant Professors comprised of 16% (the remaining 3% were non-targeted faculty members, e.g., graduate students or teaching academic staff), a ratio that reflects the overall trends in recent analysis on contingent faculty in R1 institutions and history faculty careers. Nearly three-quarters of the sample set were at least 45 years old at the time of the survey. The age ranges generally correspond with faculty rank and year of degree completion, with older respondents tending to hold higher rank and have an earlier degree completion date. Nearly 80% of the respondents earned their degree after 1990, with a mean year of 1999. Approximately half of the full professors earned their Ph.D. between 1980 and 1999, while most of the associate professors earned their degree between 2000 and 2009. Over 90% of the assistant professors who answered this question earned their doctorate between 2010 and 2019.

Previous research and commentary suggested possible links between historians’ age, career stage, and generational factors in shaping their approach to digital archival collections and related digital technologies. Yet there does not appear to be a consensus on how and to what degree these factors influence use, citation, or other engagement with digital archival collections among historians. As Alex H. Poole suggests, previous research in this vein demonstrates that “neither technophobia nor technophilia was the strict preserve of any age group.” Similarly, we do not make any assertive conclusions about whether age, rank, or generational factors affect the perceptions or use of digital archival collections. However, we occasionally note differences in the use, citation, and perceptions of digital archival materials by different faculty positions to illustrate some of the variation within a relatively homogenous sample set.

**Use and Citation of Physical and Digital Archival Collections**

Before collecting data about the participants’ perspectives of digital archival collections, we wanted to verify their general familiarity with archival materials. The responses indicated that tenured or tenure-track participants have significant experience using physical archival materials, with over 87% of reporting that they have cited these in one or more publications (see Table 1). Almost 97% reported using digital archival collections but only 88% cited them in one or more publications. The comparatively higher reported use of digital archival collections suggests that they serve purposes other than as primary source evidence in research, including for teaching and professional training—a common practice confirmed by other data from our study and detailed in previous research.

---


29 Poole, “Archival Divides and Foreign Countries?” 405.

Table 1: The Number and Percentage of Assistant, Associate, and Full Professors Who Have, Have Not, or Were Not Sure if They Have Cited Physical Archival Materials, Used, or Cited Digital Archival Collections

<table>
<thead>
<tr>
<th>Title</th>
<th>Cited Physical</th>
<th>Used Digital</th>
<th>Cited Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>547 (94.6%)</td>
<td>559 (96.9%)</td>
<td>504 (88.1%)</td>
</tr>
<tr>
<td>No</td>
<td>13 (2.2%)</td>
<td>12 (2.1%)</td>
<td>36 (6.3%)</td>
</tr>
<tr>
<td>Not Sure</td>
<td>18 (3.1%)</td>
<td>6 (1.9%)</td>
<td>32 (5.6%)</td>
</tr>
</tbody>
</table>

In Table 2, we show the percentage of respondents based on their position title. The numbers demonstrate little variation among Assistant, Associate, and Full Professors between the three comparative categories. Regardless of title, the respondents cited physical, used digital, and cited digital at about the same rate.

Table 2: Comparison of Assistant, Associate, and Full Professors Who Have Cited Physical Archival Materials to those Who Used and Cited Digital Archival Materials According to Position Title

<table>
<thead>
<tr>
<th>Title</th>
<th>Cited Physical</th>
<th>Used Digital</th>
<th>Cited Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant</td>
<td>95 (97.9%)</td>
<td>96 (99.0%)</td>
<td>84 (87.5%)</td>
</tr>
<tr>
<td>Associate</td>
<td>199 (94.8%)</td>
<td>204 (97.1%)</td>
<td>180 (86.1%)</td>
</tr>
<tr>
<td>Full</td>
<td>253 (93.4%)</td>
<td>259 (95.9%)</td>
<td>240 (89.9%)</td>
</tr>
</tbody>
</table>

About 78% of all respondents said that they have used digital archival collections more in the past 5 years, with about 22% indicating that their use has stayed about the same. As noted in Figure 1, approximately 65% (n=379) of our respondents indicated that their use and citation of digital archival collections has increased primarily because of increased availability in their research areas. This should be of little surprise given the efforts that archivists and librarians have made to digitize archival content and make available digital collections in recent years.31

**Figure 1: Reasons why participants’ citations of digital archival collections has increased in the past 5 years**

To our participants, digital archival collections are valuable for the very reasons that archivists intend them to be valuable—expanded access to archival content. As seen in Figure 2, remote access to primary sources was the most frequently cited major benefit of digital archival objects, while other benefits associated with convenience (cost savings, source location, and project planning) were also frequently selected.

**Figure 2: Major benefits of digital archival objects as indicated by participants**
As several participants noted, with digital objects researchers may download or save the images to their personal computers, thereby creating their own digital archives. This allows them to recall, review, and compare materials at a much faster rate. R168 wrote that one of the benefits of these materials is the “access to quality images for my own later reference; not having to transcribe items while IN the archive to maximize research time on-site.” R265 made a similar comment but also noted that digital objects make it easier to compare resources from multiple collections: “Saves time. Even when I am near a major research [repository] that holds the originals I need to consult, it is often more time efficient to call up the online version rather than walk to the library. Also, when looking at archives in the original in a research library I can simultaneously look at sources in other archives that are available online - important for cross referencing and checking.”

Examining the benefits according to academic rank (Table 3), we see that some of our participants remarked that digital archival collections are a useful way to begin to plan research projects. As R605 commented, reviewing digital materials is a “fun way of exploring ideas I am developing—a great way to test the waters on a topic.” R53 does not necessarily use the materials to begin a research project but to help complete it: “It’s great for filling gaps left from traditional research. Not a substitute for archival research.”

Table 3: Perceived benefits of digital archival collections per academic rank

<table>
<thead>
<tr>
<th></th>
<th>Remote access to primary sources</th>
<th>Cost savings</th>
<th>Ability to better plan research projects</th>
<th>Information about the object (i.e., its metadata)</th>
<th>Quality of objects</th>
<th>Ability to search and locate sources</th>
<th>User interface that facilitates access</th>
<th>Other (please explain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant Professor</td>
<td>85.6%</td>
<td>68.0%</td>
<td>48.5%</td>
<td>25.8%</td>
<td>6.2%</td>
<td>66.0%</td>
<td>37.1%</td>
<td>13.4%</td>
</tr>
<tr>
<td>Associate Professor</td>
<td>85.2%</td>
<td>63.8%</td>
<td>47.6%</td>
<td>27.1%</td>
<td>16.2%</td>
<td>63.3%</td>
<td>29.0%</td>
<td>12.4%</td>
</tr>
<tr>
<td>Full Professor</td>
<td>88.9%</td>
<td>68.3%</td>
<td>37.3%</td>
<td>20.3%</td>
<td>7.7%</td>
<td>64.6%</td>
<td>29.2%</td>
<td>8.1%</td>
</tr>
</tbody>
</table>

In terms of the data, Full Professors did not select this option as frequently as Assistant or Associate Professors. One possible reason why Full Professors do not rely as much on digital collections at the beginning of the research process is because they may be too accustomed to other research methods and habits. In other words, when they begin a project, they may not consider exploring digital content. As shown in Table 4, younger participants selected Ability to better plan research projects at a slightly higher rate than all the other age categories, implying that younger researchers may be more accustomed to examining digital archival collections at the beginning of their research.

Table 4: Benefits of digital archival collections per age range

<table>
<thead>
<tr>
<th></th>
<th>Remote access to primary sources</th>
<th>Cost savings</th>
<th>Ability to better plan research projects</th>
<th>Information about the object</th>
<th>Quality of objects</th>
<th>Ability to search and locate sources</th>
<th>User interface that facilitate s access</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 - 34</td>
<td>24.3%</td>
<td>18.7%</td>
<td>14.0%</td>
<td>7.5%</td>
<td>0.9%</td>
<td>17.8%</td>
<td>11.2%</td>
<td>5.6%</td>
</tr>
<tr>
<td>35 - 44</td>
<td>25.0%</td>
<td>19.9%</td>
<td>14.6%</td>
<td>7.2%</td>
<td>2.5%</td>
<td>19.0%</td>
<td>8.3%</td>
<td>3.5%</td>
</tr>
<tr>
<td>45 - 54</td>
<td>24.6%</td>
<td>20.0%</td>
<td>13.5%</td>
<td>7.5%</td>
<td>3.7%</td>
<td>18.7%</td>
<td>9.1%</td>
<td>2.9%</td>
</tr>
</tbody>
</table>
Limitations of the Use & Citation of Digital Archival Collections

Despite the benefits of digital archival collections, they are far from perfect resources. As indicated in the literature review, prior research suggests that academic historians have not always openly embraced the use and value of digital archival materials, especially in comparison to physical archival collections. We asked participants if they trust digital archival objects more, less, or about the same if they come from a reliable source, such as a university archives, compared to physical archival materials. Over 80% of the respondents indicated they trust these objects about the same as their physical counterparts, while about 16% indicated that they trust them less. Slightly less than 2% indicated they trust them more. Of the respondents who indicated they trust digital objects less, senior faculty (i.e., Associate and Full Professors) selected this response at a rate more than double their junior faculty counterparts—7% of Assistant Professors compared to 16% of Associate Professors and 19% of Full Professors. While the data from the survey is not sufficient to explain this difference, it may suggest that younger faculty have a more favorable view of digital objects than older faculty. However, this may also indicate that tenured faculty are more cognizant of the limitations of these materials because they have had more experience with them. But what are these limitations? What affects a user’s trust of digital content? Based on the responses from our survey, we contend that there are three overlapping factors that help address both questions: access, selection, and context.

Access

Although most of the respondents (approximately 57%) indicated that there is either a good or exceptional volume of material for them to use and cite in their research areas, not all research areas have widely available digital archival collections. Geographical and political issues may inhibit the access, availability, and quality of digital content, especially for historians whose research areas include African, Asian, or Eastern European history. R354 summarized such problems: “For historians, access to archives - both digital and physical - varies considerably according to geographic field. Even the physical archives I access as an Africanist are far skimpier and difficult to access than physical archives for Americanists, for example.” But it is not just the quality of materials that hinders the discovery process; other researchers mentioned that poor cataloging, metadata, or even digital interfaces constrain the usability of digital archival objects. R232 wrote that “Digitization is very slow to come to my field (South Asian History). Often the work is not fully overseen by knowledgeable archivists or catalogers familiar with non-English language works. Lot of problems in the search tools to find digital sources.”

The use and citation of digital archival collections is also limited by participants who rely on resources provided by vendors or services that require payment to gain access to the digital content. For example, R90 wrote that their “biggest concern I suppose is that archive[s] end up behind a paywall and then researchers in the global south don’t get subscriptions to the databases that package and sell digital access.” R211 echoed this sentiment: “I have recently noticed a lot

<table>
<thead>
<tr>
<th>55 - 64</th>
<th>65 - 74</th>
<th>75 - 84</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.4%</td>
<td>28.6%</td>
<td>34.5%</td>
</tr>
<tr>
<td>20.2%</td>
<td>19.2%</td>
<td>20.7%</td>
</tr>
<tr>
<td>12.1%</td>
<td>9.4%</td>
<td>10.3%</td>
</tr>
<tr>
<td>7.7%</td>
<td>4.7%</td>
<td>6.9%</td>
</tr>
<tr>
<td>2.9%</td>
<td>4.0%</td>
<td>3.4%</td>
</tr>
<tr>
<td>18.3%</td>
<td>22.1%</td>
<td>20.7%</td>
</tr>
<tr>
<td>9.1%</td>
<td>9.8%</td>
<td>3.4%</td>
</tr>
<tr>
<td>3.3%</td>
<td>2.2%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

32 We did not ask for the respondents’ areas of research, but some participants volunteered this information.
more imbalance in the commercial archives that my university subscribes to, in regards the kinds of material that we tend to purchase (focusing on white and/or Asian history). Some of the archives we can access via ProQuest or other sources often have very limited content related to African American and other minority history.” R211 also mentioned the larger implications of these restrictions: “The gatekeeping aspects and potential for implicit bias in the major providers of archival content is concerning for me as it means we are potentially limiting access to important sources for ourselves and our students.”

We did not anticipate these responses because none of the participants in our pilot study mentioned using vendor-based digital archival collections. The responses we received from the survey imply that it may be a growing concern among academic historians. More importantly, this issue raises the concern of the discoverability of archival digital content that is publicly available. Several of our participants expressed frustration about the difficulties they have in locating digital sources. For example, R281 wrote that “Because documents—sometimes from the same archive—are often scattered across several digital collections, it is often hard to find digital sources. It can also be cumbersome to repeat searches in multiple digital databases.” R413 remarked that it is easier to locate sources using physical archives than it is online: “Digital source location is not always well organized in my field; that is, there is no single easy way to find them. Early printed books are easier to find than digitized manuscripts.” And R304 and R440 noted the lack of a centralized portal for digitization projects and collection: “The one thing that I find most difficult is learning about and/or finding new digital archival collections, as I don’t know of a centralized portal to learn about ongoing digitization work” (R304) and “In my field, we need a hub for identifying all archives, and tracking as well those that are maintained” (R440).

Selection

The second issue that affects the trust of digital archival collections is the selection of what materials are digitized and placed online. Academic historians understand digital archival collections that appear online seldom, if ever, fully represent the contents of the physical collection. In fact, several of our participants expressed some frustrations about the selection process that archivists use to determine what should be digitized. R67 noted that the historical profession “is at the mercy of which archives have promoted digital access (and in those archives, which materials have been digitized).” R621 offered one of the most critical responses about digital archival materials that we received: “Partially digitized collections may be the worst because the criteria of what to digitized and what not to are often opaque, unexplained, or even unstated.” The resulting fragmentation creates a paradoxical situation for users: they

---

34 <Citation Anonymized>.

35 Mark Sandle observed that historians “need to be wary of this idea that there is an inherent authenticity that comes with handling directly the traces of the past. Archivists select, sift and discard. We need to remember that all of our encounters with the lingering remnants of the past have to be appraised critically, carefully and appropriately” (Mark Sandle, “Studying the Past in the Digital Age: From Tourist to Explorer,” in History in the Digital Age, ed. Toni Weller (New York: Taylor & Francis Group, 2012), 136).

36 Ogilvie, “Scientific Archives in the Age of Digitization,” 83. Ogilvie refers to this as “virtual dismemberment,” where partial the presence of digital archives can obscure the overall structure of the corresponding physical archives.
potentially allow for more extensive and greater equity in access, but the selected digital surrogates may also influence what research can be accomplished online. As R62 remarked, digitized content “pushes historiography to sexier topics - genocide, race, etc., since the easier access to materials inevitably makes them more widely used.” R67 observed that digital archival collections increase “access but (potentially) decreases the breadth of material looked at” and by doing so, they “can eliminate the eureka moment when one discovers material that hasn’t been cataloged well or has been mis-described in an inventory.”

Based on these comments, archivists need to be aware that selective digital content may distort the original context of the materials placed online, which may offset their value or affect how users interpret them. According to several participants, the challenge lies in reconstructing context. For example, R220 noted that digital archival collections “don’t necessarily maintain a sense of original ‘context,’ in terms of what accompanied the individual object when it arrived in the collection. Often, that is key information for interpretation.” R191 wrote that in “many cases it is still hard to understand what documents physically are next to digitized materials (context) even if an online finding aid is available, if only selected items have been digitized.” The difficulty in assessing context may influence how trusting some users are of digital archival collections and their content because it potentially threatens the usefulness of the object, a point made by R413 who noted that “I would mark unreliable too, but in one particular way, in that they are often incomplete. Usually in my field, if something is digitized, the digitization itself is reliable.” Here, we see the issue of not trusting how the object may be interpreted because its context is not understood as opposed to not trusting the object because of questions concerning its authenticity. Therefore, what materials archivists place online directly results in the most concerning issue that our participants have of these materials: understanding their context.

Context
According to our participants, the most significant challenge to trusting digital archival collections is understanding the context of the object. As shown in Figure 3, the two most selected limitations were Knowing that something is missing or has not been scanned (i.e., contextual concerns) and Poor user interface; and only a very small percentage believe that digital archival collections are unreliable. In none of these questions were we able to discern significant differences among the respondents according to position title, age range, or career stage (see Table 5). The data points to a broad consensus within our sample group that concerns over context represent the most substantial limitation to digital archival collections. This is verified and reiterated in the open-ended responses we received from dozens of survey participants.

---

37 R315 “As a historian working on foreign subjects, the ability to access digitized sources in lieu of flying trans-Atlantic is irreplaceable. Digitization improves equity among researchers: women, poc, parents, and less-affluent can still access relevant materials and conduct research.”
Figure 3: Major limitations of digital archival collections

Table 5: Major limitations of digital archival collections per academic rank

<table>
<thead>
<tr>
<th>Knowing that something is missing or has not been scanned</th>
<th>Poor user interface</th>
<th>Poor quality of the digital archival objects displayed</th>
<th>Unable to locate digital archival objects relevant to my research</th>
<th>Too many digital archival objects to review</th>
<th>Other (please explain)</th>
<th>Digital archival objects are unreliable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant</td>
<td>24.4%</td>
<td>24.8%</td>
<td>16.2%</td>
<td>15.8%</td>
<td>9.0%</td>
<td>6.4%</td>
</tr>
<tr>
<td>Associate</td>
<td>26.1%</td>
<td>22.6%</td>
<td>14.3%</td>
<td>15.9%</td>
<td>11.3%</td>
<td>8.1%</td>
</tr>
<tr>
<td>Full</td>
<td>27.9%</td>
<td>23.7%</td>
<td>15.4%</td>
<td>13.1%</td>
<td>9.2%</td>
<td>7.1%</td>
</tr>
</tbody>
</table>

Over a third of these responses made some mention of needing to understand context. For example, R220 wrote that digital does not “necessarily maintain a sense of original ‘context,’ in terms of what accompanied the individual object when it arrived in the collection. Often, that is key information for interpretation.” R355 echoed this point by saying that digital objects are often “extracted from a larger collection. In the past, I have come across other sources I didn’t know existed by searching through a larger series of documents. This is a real problem.” R47 noted the paradoxical situation that digital archival collections create: “Although digital archives often expand research capacity, they can also narrow research in that they are necessarily a limited, non-random sample of available material.” The selection process is what may continue to drive researchers to the archives, as R443 wrote: “Sometimes digitized objects hint at what other, non-digitized objects, are in a collection […] And, I add, it is important to get local context, so as much as I like digital archives, I think visiting the research location is important, and particularly for those of us doing international work.” R432 likened viewing archives to browsing the stacks in a library: “I think of it the same way I think of the library shelves; I need to see what is around it to get a sense of it is representative of a pattern or a tendency or
potentially anomalous with other similar objects.” And R193 commented on the idea of what is lost by not visiting an archives: “Too many colleagues search digitalized periodicals using key words, which dramatically undercuts both contextual understanding or unexpected discoveries.” In other words, searching online may not be the perfect substitute for browsing through physical collections in the archives.

As in previous studies of academic historians and digital archival collections, the importance of context emerges in our data, but in way that suggests it outweighs other concerns related to reliability, trustworthiness, materiality, and system usability. Our academic historians do not seem to question the trustworthiness of most digital surrogates; they are more concerned with ascertaining information about the digital object in its original context rather than worrying if the object is a forgery. For them, trust stems from the notion that individual objects do not exist in a vacuum and cannot be accurately understood without knowing how objects relate to adjacent materials in the collection, otherwise known as the archival bond. Knowing the original context of the object represented by the digital surrogate is vital to understanding the historical phenomena associated with the documentation, but digital archival collections often lack relevant contextual information that might provide further clarity. Respondents repeatedly emphasized this drawback.

Typically, concerns about context implies a larger concern about the trustworthiness of digital archival collections and objects. Yet, for our respondents, questions over trustworthiness do not appear to inhibit their use or the decision to cite them in a publication. In fact, as seen in Figure 2 (above), the respondents selected Increased trust in digital archival collections and objects the fewest number of times. Data from three other survey questions also support this finding.

First, we asked if respondents still need to see the original object prior to citing it in a publication. Nearly half selected No, and the other half chose It Depends, with only a handful indicating Yes (see Table 6). We contend that if academic historians have an overall distrust of these objects, then there would be a higher percentage responding affirmatively to this question.

Table 6: Participants’ need to see the original object prior to citing it per academic rank

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Depends on the object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant</td>
<td>2 (2%)</td>
<td>36 (42%)</td>
<td>47 (55%)</td>
</tr>
<tr>
<td>Associate</td>
<td>5 (3%)</td>
<td>66 (36%)</td>
<td>110 (61%)</td>
</tr>
<tr>
<td>Full Professor</td>
<td>10 (4%)</td>
<td>120 (49%)</td>
<td>114 (47%)</td>
</tr>
<tr>
<td>Overall</td>
<td>17 (3%)</td>
<td>222 (43%)</td>
<td>271 (53%)</td>
</tr>
</tbody>
</table>

This is not to say that the original object is unimportant. According to some of our historians, the original is important to ensure that they can verify that the digital surrogate is free from blemishes or errors that may have occurred during the scanning process and so they may accurately assess the context of the object. R132 wrote that “Sometimes there are markings you may only notice in person, or may only be legible in person; sometimes I may need to know what else is in a collection that hasn’t been digitized in order to have the entire context for the digital object.” R78 made a similar statement by saying that they prefer the original to “usually

to read handwriting more closely, or to see other documents in context not included in digital collection.” And R617 stated that these materials “are wonderful, particularly for students who don’t have the opportunity to travel to archives. But it doesn’t make up for working in the archival collections, themselves, because one is dependent upon the selections made by those who have digitized collections, and by the way in which they have envisioned those collections’ utility.” Once again, we see that context plays a more important role in evaluating these materials than does assessing their trustworthiness.

Second, for the participants who chose *Yes* or *It Depends* on whether they need to see the original object before citing it in a publication, we asked why they need to see it. Our rationale for this line of questioning is that some historians have obsessed with the notion of the original, with at least one scholar going so far to say it is a “fetish.”* As shown in Figure 4, our respondents selected the option that addresses understanding context at over three times the rate than the second most selected option—*Verify the authenticity of the object*. Although we did not specifically ask how our respondents define the concept of trustworthiness and if and how they make any distinctions between the concept of authenticity and reliability, some participants indicated that once they see that a digital archival collection is hosted by an institution that they consider to be reliable, they are no longer concerned with the trustworthiness of the collection. For example, in response to a question about how they know that they are seeing the object in the original context, R516 wrote: “Institutional trust. The digital archives I use most frequently were assembled by the French national library and national archives. This means I can trust the integrity of their contents in ways I, of course, would not be able to trust random scans of things put up by amateurs.”

*Figure 4: Why do you want to see the original object?*

As shown in Table 7, full professors seem to not trust digital objects at a slightly higher rate than associate or assistant professors demonstrated by a higher percentage of them wanting to see the original so they can verify the authenticity of the digital surrogate. Assistant professors

---

prefer to see the original to better understand the context of the object at a higher rate than senior faculty.

Table 7: The reasons participants need to see the original object per academic rank

<table>
<thead>
<tr>
<th></th>
<th>Verify the authenticity of the object</th>
<th>Verify the quality of the object</th>
<th>Better understand the context of the object</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant Professor</td>
<td>7 (12%)</td>
<td>4 (7%)</td>
<td>40 (68%)</td>
<td>8 (14%)</td>
</tr>
<tr>
<td>Associate Professor</td>
<td>30 (18%)</td>
<td>10 (6%)</td>
<td>98 (59%)</td>
<td>27 (16%)</td>
</tr>
<tr>
<td>Full Professor</td>
<td>39 (21%)</td>
<td>26 (14%)</td>
<td>95 (52%)</td>
<td>22 (12%)</td>
</tr>
</tbody>
</table>

The majority of the 57 “Other” responses overlap with the other options, with most of them explaining that the original object allows them to review or verify its content and better understand its materiality. R88 wrote that the original allows them to “see what else the document contains. One problem is seeing the whole thing, which gives clues, context or additional information that are not apparent on the digital site.” R295 noted that “Sometimes details on a digital image are difficult to discern, and seeing something in person may clarify (like, a handwritten note on the back of an official document may be blurry in the digital image).” R559 made a similar statement by observing that the originals are needed to reveal “details that might not appear on a scan, e.g., watermarks, quality of paper, whether a mark is a pen stroke or a wormhole, etc.” And R576 lamented that “Digitizations are rarely of sufficiently high quality with adequate contextual data to be entirely certain of the object. That is the issue is the unknown reliability of the digitization and its tacit characterization of the nature of the object.”

Our data reveals that there continues to be some distrust in some of the content that academic historians view online. These users are wary that the digital version is not the complete and accurate version of the original object. Some of this trust may result from one’s research area and the notion that the digitization process may not be equally done at all institutions around the globe, but most of the distrust comes from the inability of users to not be able to discern the context of online objects.

The participant observations add another layer of complexity to the concept of context. The Society of American Archivists (SAA) definition of context includes “the circumstances that a user may bring to a document that influences that user’s understanding of the document.” Our study supports the notion that the understanding of context is shaped by personal notions of trustworthiness, authenticity, and reliability—interrelated concepts that may hold different meanings to different creators and users of digital archival collections.

Conclusion

Overall, this study shows that the use and citation of digital archival collections has increased among the respondents in past five years and there is no reason to believe that this trend will change anytime soon. As researchers have fewer opportunities to visit archives in person due to health and safety concerns, restricted budgets, or other considerations, archivists must keep in mind that users seek out these materials for various reasons. While the focus of this study was on research use and citation, respondents indicated multiple uses related to teaching,

---

professional education, and accumulating subject knowledge. Thus, archivists should approach the development of digital archival collections with consideration of these uses in addition to the potential long-term consequences for historiography. As historians use and cite digital archival collections more frequently, archivists must continue to be aware of their impact on the historical record and how their decisions influence its interpretation by historians.

Digital archival collections have helped our respondents advance their research agendas because of how easily they may be accessed. As R25 noted, “digital sources are not going away. Thankfully, I look forward to ever more collections being available online.” R68 remarked that these materials are a “vast improvement for my research since I began doing it in the 1960s. I cannot imagine doing [my research] without this tool.” R103 echoed these sentiments: “I feel blessed to be able to access so many archival materials digitally. It is a great convenience—a time saver and a money saver.” And finally, R90 wrote: “I love digital archives. Any scanning is better than no scanning... even if the archive is thin, it allows a minimum reconnaissance and that is one less object I have to view on site. My research is expensive, so I welcome any opportunity to ease the access.”

Even with their positive features, digital archival collections are not without their limitations. As one scholar observed “historians have always had to address logistical problems in doing historical research. The digital age has solved some problems of access, but created others.” This study suggests that a paradoxical situation frequently unfolds, whereby the ease of access potentially shapes historical understanding and interpretation in ways that are difficult (if not impossible) to anticipate and not always desirable. The effects of archival selection and limited digital content permeate throughout the survey and pose challenges to how users interact with and interpret online materials. As R3 succinctly noted: “Access should never be confused with understanding.”

Not only do archivists face the pressures to digitize more content but they also need to find ways to make this content visible to users. Additional research needs to assess how users, such as academic historians, locate these sources. It is unlikely that there will ever be a single system that provides access to digital archival collections among all archival institutions, but there are several projects that are creating centralized portals for thematically-linked collections. Hypothetically speaking, even if archivists were to resolve the larger discoverability issue using a search engine to locate digital resources, as R79 observed, “While I am comfortable using digital archival collections, I still find finding aids more useful than digital search tools for determining whether I’ve located and examined the relevant files. I find the digital search engine version of collection display often obscures relevant documents that I would have found using a traditional finding aid.”

Careful attention to the presentation of the content and how users interact with it is important for the design of digital archival collections. Archivists must also assess what information should be associated with these materials so that academic historians and other users may more accurately discern context to interpret the materials they encounter. Moreover,

---

42 Sandle, 133.
43 Sinn and Soares, 1804-1805.
44 There are a couple national portals that serve as guides for digital content, most portals remain regional or based on specific topics. NCPH: https://ncph.org/digital-public-history-directory/. See also: Amy Crawford, “A Massive New Database Will Connect Billions of Historic Records to Tell the Full Story of American Slavery,” Smithsonian 50, no. 9 (2020): 1.
archivists must also consider new ways to be more transparent with their selection processes that may increase the degree of trust that users have of digital content. As historians increasingly incorporate digital archival collections into their regular workflows, archivists must be mindful of how their decisions to privilege some records over others impacts the production of history.
Appendix A
Q2.1 Have you ever used a digital archival collection for research purposes?

"Use" means that you have referred to a digital archival collection or a digital object within a digital collection and it assisted with your research but you did not cite either in a publication.

"Digital archival collection" means a group of digitized material typically found in an archives, such as photographs, text documents, manuscripts, artwork, audio files, etc.

For the remainder of this survey, we use the phrase "digital archival object" with the understanding that it refers to the content in digital archival collections.

- Yes
- No
- Not Sure

Q2.2 Have you ever cited a digital archival object in a publication, such as an article, monograph, chapter, etc.?

- Yes
- No
- Not Sure
Q2.3 Please indicate the reason(s) you have not cited a digital archival object in any of your publications. Please select all that apply.

- Cannot locate any relevant digital archival collections in area(s) of research
- Poor quality of digital archival collections or objects
- Prefer to cite the original (i.e., physical) materials
- Do not trust the digital object as being authentic or reliable
- Digital object lacks sufficient contextual information to be used as a reliable resource
- Scholarly community (i.e., publishers, peers, etc.) discourages the use of digital archival objects
- Other (please explain): __________

Q3.1 Have you ever cited one or more physical archival materials in a publication?

- Yes
- No
- Not Sure
Q3.2 In how many publications (i.e., individual articles, books, book chapters, etc.) have you cited physical archival materials?

- 1-10
- 11-15
- 16-20
- 21-25
- 26-30
- 31-35
- 36-40
- 41 or more

Q4.1 What is your age?

- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older
Q4.2 What is your current position title?

- Adjunct/Lecturer
- Assistant Professor
- Associate Professor
- Distinguished Professor
- Emeritus Professor
- Full Professor
- Postdoctoral Researcher/Fellow/Scholar
- Other (please specify): __________

Q4.3 In what year did you earn your PhD?

Q4.4 Would you be willing to be contacted for a possible follow-up survey or interview about your use of digital archival collections?

- Yes
- No

Q4.5 Thank you for your willingness to further discuss your experiences with digital archival collections with the research team. Please include your name & email address so we may contact you.

- Name ____________________________________________
- Email ____________________________________________
Q5.1 How confident are you that there are digital archival collections in your primary area of research?

- Extremely confident
- Moderately confident
- Slightly confident
- Neither confident nor not confident
- Not confident at all

Q5.2 When engaged in a research project, on average, how often do you search for and interact with digital archival collections?

- Multiple times per day
- Multiple times per week
- Multiple times per month
- Monthly
- Infrequently

Q5.3 How abundant are relevant digital archival objects in your primary area of research?

- Exceptional volume (can always find resources)
- Good volume (can often find resources)
- Fair volume (can sometimes find resources)
- Limited volume (can rarely find resources)
- Poor volume (can never find resources)
- Do not know or cannot estimate

Q5.4 In how many publications (i.e., individual articles, books, book chapters, etc.) have you cited digital archival collections or objects?
Q5.5 Once you see a digital archival object online, do you typically want to see it in person before you are willing to cite it in your research?

- Yes
- No
- Depends on the object

Q5.6 Why do you typically need to see the original archival object before you are willing to cite it in your research? Please select all that apply.

- [ ] Verify the authenticity of the object
- [ ] Verify the quality of the object
- [ ] Better understand the context of the object
- [ ] Other (please specify): __________

Q5.7 When looking at a digitized object online, as part of a digital collection, how important is it to you that you understand the object in its "original" context (i.e., what you see online is what
you would expect to find in the corresponding physical collection if you were accessing it at the holding institution in person)?

- Extremely important
- Very important
- Moderately important
- Slightly important
- Not at all important

Q5.8 How do you know that you are seeing the object in the "original" context? Select all that apply.

- Verify that the object is accounted for in the collection’s finding aid (if available)
- Look for context clues on the object, such as serialization or numbering, changes in visual features, changes in subject matter, etc.
- Access the original object at the archives
- Contact the archives and inquire about the original object
- I don’t, but I’m okay with that
- Other (please explain): __________
Q6.1 What do you see as the major benefits of digital archival objects?

☐ Remote access to primary sources

☐ Cost savings (i.e., not needing to travel)

☐ Ability to better plan research projects

☐ Information about the object (i.e., its metadata)

☐ Quality of objects

☐ Ability to search and locate sources

☐ User interface that facilitates access

☐ Other (please explain): __________

Q6.2 What do you see as the major limitations of digital archival objects?

☐ Too many digital archival objects to review

☐ Knowing that something is missing or has not been scanned

☐ Digital archival objects are unreliable

☐ Unable to locate digital archival objects relevant to my research

☐ Poor quality of the digital archival objects displayed

☐ Poor user interface

☐ Other (please explain): __________
Q6.3 Assuming they come from a reliable source, such as a university archives, do you trust digital objects more or less than physical archival materials?

- More
- Less
- I trust them about the same

Q6.4 Within the past 5 years, has your use of digital archival collections increased, decreased, or stayed about the same? Please base your response on your research habits prior to the COVID-19 pandemic.

- Increased
- Decreased
- Stayed about the same

Q6.5 Why have you cited digital archival collections more in the past 5 years? Select all that apply.

- Increased trust in digital archival collections and their objects
- More digital collections and/or objects have become available in field of research
- Decrease in funding to travel to archives to see materials in person
- Decrease in time to travel to archives to see materials in person
- More comfortable locating digital archival collections
- Other (please explain): __________
Q6.6 Please indicate why you have cited fewer digital archival collections in the past 5 years? Select all that apply.

- [ ] Have conducted fewer research projects that require these types of resources
- [ ] Need them primarily for teaching, not publishing
- [ ] Too time consuming of a process to locate digital objects
- [ ] Lack of new digital objects in my research area
- [ ] Encountered too many objects of poor quality
- [ ] Context is never sufficient to fully understand the digital object
- [ ] Do not sufficiently trust the digital objects (even if they come from a reliable source, such as a university archives)
- [ ] Other (please explain): __________

Q7.1 How important is the usability of the interface when you search for and want to use digital archival collections for your research?

- [ ] Extremely important
- [ ] Very important
- [ ] Moderately important
- [ ] Slightly important
- [ ] Not at all important

Q8.1 Do you have any final comments or observations about your use of digital archival collections? If so, please provide in the following space.