


Sowing the Seeds for More Usable Web Archives: A Usability Study of Archive-It

Samantha Abrams, Alexis Antracoli, Rachel Appel,
Celia Caust-Ellenbogen, Sarah Denison,
Sumitra Duncan, Stefanie Ramsay

ABSTRACT

In 2017, seven members of the Archive-It Mid-Atlantic Users Group (AITMA) conducted a study of fourteen participants representative of their stakeholder populations to assess the usability of Archive-It, a web archiving subscription service of the Internet Archive. While Archive-It is the most widely used tool for web archiving, little is known about how users interact with the service. This study investigated what users expect from web archives, a distinct form of archival materials. End-user participants executed four search tasks using the public Archive-It interface and the Wayback Machine to access archived information on websites from the facilitators' own harvested collections and provided feedback about their experiences. The tasks were designed to have straightforward outcomes (completed or not completed), and the facilitators took notes on the participants' behavior and commentary during the sessions. Overall, participants reported mildly positive impressions of the Archive-It public user interface based on their sessions. The study identified several key areas of improvement for the Archive-It service pertaining to metadata options, terminology display, indexing of dates, and the site's search box.

© Samantha Abrams, Alexis Antracoli, Rachel Appel, Celia Caust-Ellenbogen, Sarah Denison, Sumitra Duncan, Stefanie Ramsay. 

KEY WORDS

Web archives, Usability, Usability testing, Archive-It, User interfaces,
Archival discovery systems, User research

A popular misconception holds that once something is on the Internet, it will be there forever, but many information professionals know this is not the case. To ensure that information of enduring value is preserved and accessible in the future, archivists increasingly rely on web archiving tools to take snapshots of websites at particular moments. As a type of archival material, web archives are increasing in importance and use. Capturing, preserving, and accessing them require specialized tools.

According to the National Digital Stewardship Alliance 2017 Survey, “Web Archiving in the United States,” Archive-It is by far the most popular web archiving tool in use.¹ Surprisingly, little research has focused on web archiving tools, with an even smaller subset of existing literature addressing the usability of these services. The Archive-It Mid-Atlantic Users Group (AITMA) formed a working group in 2017 to assess the usability of Archive-It with regard to end-user experience in searching for and accessing information archived from the Internet. This working group comprised seven archivists and librarians from Delaware Public Archives, the Frick Art Reference Library, Ivy Plus Libraries, Princeton University, Swarthmore College, and Temple University.

By assessing participant engagement with the field’s most used web archiving service, we are attempting to bridge the gap between web resources and understanding how they are used. Though several institutions around the world have completed surveys and projects concerning their own web collections, none have explored the relationship between Archive-It and its users. Through our research, we hope to shed light on the way users interact with a major service that provides access to archived web collections.

After an initial review of existing literature related to web archiving, the working group connected with members of their institutional networks to identify fourteen participants representing major stakeholder groups: faculty, staff, students, librarians/archivists, and members of the general public. These participants were asked to execute tasks designed to yield a manageable and easily quantifiable data set. Testers took notes on participants’ behaviors and comments made while completing the assigned tasks. The working group adopted a hybrid quantitative and qualitative approach to analyzing this data and developed a codebook consisting of twenty-four prevailing themes in the data as an aid to analysis.

This study yielded information that suggests future directions for the Archive-It service, but is far from conclusive. Overall, participants performed well, completed most tasks, and rated their general impression of the Archive-It end-user platform as 6.5 out of 10. Their moments of confusion or frustration point to a few desirable improvements, as detailed in the Recommendations section. Still, more testing in web archiving usability is

needed, and AITMA hopes that this project can serve as a model for similar studies.

This article will provide background on Archive-It, review existing literature on related topics, outline the methodology and scope of the study, describe the testing and analysis processes, identify concerns and biases, and provide recommendations based on our findings.

About Archive-It

Created in 2006 by the Internet Archive, Archive-It is a subscription web archiving service that has long been used by individuals and organizations practicing web archiving around the world. As of 2014, the service was a partner to “over 400 . . . organizations in forty-eight . . . states and sixteen countries worldwide” made up of “college and university libraries; state archives, libraries, and historical societies; federal institutions and NGOs; museums and art libraries; and public libraries, cities, and counties.”² Archive-It differs from the Wayback Machine, a public archives of websites run and maintained by Internet Archive staff, in that Archive-It allows subscribing institutions to curate collections of websites relevant to their institutions and/or aligned with their collection development practices. Websites that are captured and preserved exist as another form of archival material, equal in importance to traditional formats. They are considered part of each institution’s archival holdings and may be used in much the same way as analog or digitized archival collections.

Websites are captured (or “crawled”) with software designed by the Internet Archive, then saved on Archive-It servers. Archive-It offers a back-end interface for a subscribing institution’s archival or library staff to perform data management and metadata creation. Completed crawls can be made available and accessed through an end-user interface, as Figure 1 shows. Our study addresses only this end-user interface.

Through the end-user interface, users can explore collections of websites, or “seed” URLs, maintained by the Archive-It collection institution. Users can browse and search collections, and use facets provided to filter the displayed results based on a certain field’s contents, as seen in Figure 2.

Once a user selects a specific seed, he or she is able to see how many times and on which dates that website was captured by Archive-It software, as shown in Figure 3.

After selecting a date, users are shown the archived website view, made clear by an informational banner, as Figure 4 shows.

Temple University Special Collections

Archive-It Partner since: Jan, 2014
 Organization Type: Colleges & Universities
 Organization URI: <http://www.library.temple.edu/scrc>

Description: The Special Collections Research Center is the principal repository for and steward of the Libraries' rare books, manuscripts, archives, and University records. We collect, preserve, and make accessible primary resources and rare or unique materials, to stimulate, enrich, and support research, teaching, learning, and administration at Temple University. SCRC makes these resources available to a broad constituency as part of the University's engagement with the larger community of scholars and independent researchers. Many of the SCRC's collections documenting the Philadelphia region's history are found in our Philadelphia Dance Collection, Philadelphia Jewish Archives, and Urban Archives. These Archive-It collections capture the on-line presence of some of the organizations whose archives are preserved by the SCRC. Complete information about the archives is available at <http://www.library.temple.edu/scrc>.

Narrow Your Results

Subject Sort By: Count (A-Z)

- Society & Culture (4)
- Education - Pennsylvania - Philadelphia (2)
- Immigrants - Pennsylvania - Philadelphia (2)
- Social service - Pennsylvania - Philadelphia (2)
- African Americans (1)

More ▾

Creator Sort By: Count (A-Z)

- Athenaeum of Philadelphia (1)
- Azavea (1)
- Bicycle Coalition of Greater Philadelphia (1)
- Buy Local PA (1)
- Code for Philly (1)

More ▾

Publisher Sort By: Count (A-Z)

- Azavea (1)
- The City of Philadelphia Office of Innovation and Technology (1)

Date Sort By: Count (A-Z)

- 2011- (1)

Language Sort By: Count (A-Z)

Sites and collections from this organization are listed below. Narrow your results at left, or enter a search query below to find a collection, site, specific URI, or to search the text of archived webpages.

Enter search terms here

Collections Sites Search Page Text

Page 1 of 1 (4 Total Results)

Sort By: Collection Name (A-Z) | Collection Name (Z-A)

Jewish Archives

Archived since: Jan, 2014

Description: Acquired by Temple University Libraries' Special Collections Research Center in 2009, the Philadelphia Jewish Archives Collection documents the activities of the greater Philadelphia Jewish community through the personal papers of civic and spiritual leaders and the records of cultural, educational, religious, and social service organizations established and operated by and for the Jewish community. The Jewish Archives web archive captures web sites that complement the Philadelphia Jewish Archives Collection.

Subject: Society & Culture, Jews - Pennsylvania - Philadelphia, Education - Pennsylvania - Philadelphia, Immigrants - Pennsylvania - Philadelphia, Jews - Charities, Social service - Pennsylvania - Philadelphia, Synagogues, Publishers and publishing, Medical care - Pennsylvania

Coverage: Philadelphia (Pa.), Pennsylvania

Accessibility Note: The content presented here appears in its original, legacy formats, and may not be easily accessible. For Americans with Disabilities Act (ADA) accommodation, including help with reading this content, please contact scrc@temple.edu.

Collector: Special Collections Research Center, Temple University Libraries

Open Data Philly

Archived since: Nov, 2016

Description: Open Data Philly is an open data aggregate and portal that includes over 300 datasets from the City of Philadelphia as well as various non-profit organizations and for-profit organizations in the Philadelphia area. This collection was created because Temple University was awarded a

FIGURE 1. Institution landing page in the Archive-It end-user interface

Literature Review

In 2017, more than a hundred individuals participated in the National Digital Stewardship Alliance (NDSA) survey “Web Archiving in the United States.” The survey, also administered in 2011, 2013, and 2016, was developed to “better understand the landscape of web archiving in the United States by investigating the organizations involved, the tools and services being used, access and discovery services being provided, and overall policies related to web archiving programs.”³ “Web Archiving in the United States” not only interrogates the field of web archiving as it stands, but acts as a barometer and “tracks the evolution of the field over the last few years, and points to future opportunities and

Title: Swarthmore College website
URL: <http://swarthmore.edu>
Description: The official website of Swarthmore College. Includes college-related news and information on a wide range of College activities, including admissions, academics and athletics. Captured 50 times between Mar 7, 2011 and Apr 19, 2018
Videos: 820 Videos Captured
Subject: Swarthmore College
Group: Swarthmore
Creator: Swarthmore College
Publisher: Swarthmore College
Language: English
Coverage: Swarthmore (Pa.)
Date: 2011, 2012, 2013, 2014, 2015, 2016, 2017
Rights: This Item is protected by copyright and/or related rights. You are free to use this Item 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).
Identifier: 1376689
Collector: Swarthmore College

FIGURE 2. Seed metadata view in Archive-It

Found 60 Captures between Feb 3, 2006 - Apr 30, 2018						
2009	2010	2011	2012	2013	2014	2015
4 pages	5 pages	5 pages	4 pages	5 pages	4 pages	2 pages
Jan 25, 2009	Jan 15, 2010 *	Apr 27, 2011 *	Jan 3, 2012	Jan 6, 2013 *	May 17, 2014 *	Apr 30, 2015 *
May 18, 2009 *	Feb 18, 2010	Apr 30, 2011	Aug 14, 2012 *	Mar 7, 2013 *	Aug 1, 2014 *	Oct 13, 2015 *
Jul 28, 2009 *	Mar 28, 2010 *	Jun 2, 2011	Oct 28, 2012 *	May 30, 2013	Nov 10, 2014 *	
Aug 28, 2009	Dec 5, 2010 *	Jul 5, 2011	Nov 12, 2012 *	Sep 29, 2013 *	Dec 17, 2014 *	
	Dec 8, 2010	Dec 3, 2011 *		Dec 31, 2013		

FIGURE 3. Captures available for one website



FIGURE 4. Archived website view

developments.”⁴ The survey asked questions like “How much full-time equivalent staff time does your organization dedicate to web archiving?” and “What types of content do you have concerns about your capacity to archive?”

Most of the results from the 2017 survey changed little from previous years. One area, though, saw an increase: Archive-It use. As the authors write, “94% . . . of institutions . . . capture content with Archive-It” and “Archive-It has been the favored external service provider since NDSA began recording these statistics in 2011, when 72% of responding institutions were using it.”⁵

Archive-It’s high use makes sense, as the subscription-based web archiving service “helps organizations to harvest, build, and manage born digital collections” and “collect, catalog, and manage their collections of archived content” with round-the-clock access and full-text search capabilities without organizations having to devote significant internal resources to harvesting, preserving, and providing access to web archives.⁶ An Archive-It subscription also includes the ability to create and manage metadata, and, though the platform has no required metadata fields, it does support the use of fifteen standard Dublin Core elements and the ability to create custom metadata fields, all of which are “displayed and hyperlinked for browsing on public collection pages hosted at Archive-It.org.”⁷ In the end, it may be Archive-It’s ability to “provide the viewer a wealth of information”⁸ that makes the service so appealing to archivists. As researchers at Old Dominion University note, “Archive-It collections, in particular, are interesting because a single organization is responsible for each collection, meaning that the curation strategy for a collection is guided by humans rather than automated crawling operations.”⁹ As professionals, we can infer that Archive-It’s popularity has much to do with the ability to carry out web archiving using a platform that handles much of the technical work, while allowing archivists to provide some intuitive level of access to researchers.

And yet, despite its status as the most utilized web archiving service, professionals in the field know little about how researchers use and interact with collections in Archive-It. In fact, the field knows very little about how researchers use and interact with any web archives access tools. This contrasts with the more substantial literature on the usability of other types of archival discovery systems. Beginning in 2003, with Elizabeth Yakel’s study of an Encoded Archival Description (EAD) interface, archivists have published numerous studies about how users interact with online finding aid discovery systems. Research includes studying social interaction in finding aids, how search functionality affects discoverability, and the general usability of several different finding aid sites, either individually or comparatively.¹⁰ The usability of web archive discovery systems has been explored by archivists at national libraries with their own custom discoverability systems, and several explorations of search functionality, context, and researcher needs have been conducted. This research will

be summarized in the ensuing pages. However, to date, no study exists of how users interact with Archive-It, even though it is by far the most widely used web archiving discovery platform.

As Helen Hockx-Yu stated in 2014: “Web archiving institutions have traditionally focused their attention on developing the processes and technology for data collection” while the “use of web archives has not been a key element of the strategy.”¹¹ Others agree. As David Cruz and Daniel Gomes found in 2013, researchers and “most web users are not acquainted with web archives and accessing . . . archived web data is a significantly different user experience than accessing the live web.” They also report that “although there [are] a significant number of web archives available, only preliminary research has been done about the design of user interfaces to gain access to temporal web data.”¹² Additionally, as is mentioned in “Researcher Engagement with Web Archives: Challenges and Opportunities for Investment”: “We know some of what researchers want from web archives and we are just now beginning to know what web archives can provide, but web archives of use to researchers have not been available long enough yet to develop a comprehensive understanding of how researchers are engaging with web archives.”¹³

What do we know, then? The study conducted by Cruz and Gomes revealed that users of web archives like having the ability to search—as long as a bit of guidance accompanies that ability. “Searching historical web content was an awkward concept for most users,” they found, and “the existence of a website that provides access to pages that are no longer available on their original websites is a perception that requires technical knowledge about the functioning of the Internet that is far beyond the skills of common web users.”¹⁴ Cruz and Gomes made these observations when looking at the Portuguese Web Archive and noted that an explanation of, and context surrounding, the archive was necessary. They point out that, “unlike most live-web search engines, our web search homepage also includes a fat footer. The objective was to provide additional links to information that clarifies users about the context of web archiving and web archive search.”¹⁵ In fact, for users of the Portuguese Web Archive to be successful, Cruz and Gomes found that “search interfaces for web archives must be similar to live-web search engines to facilitate the adoption of web archives by new users” alongside ample context, including a “homepage for a web archive search service” that presents “contextual information such as examples of archived pages, that enable new users to gain awareness about what is, and the potential of searching, a web archive.”¹⁶

This request for context is a popular one. In “Web Archives for Researchers, Representations, Expectations, and Potential Uses,” the authors write that “Regarding services and promotion, the decisions and criteria that are used in the selection policy must be justified, documented, and made visible.”¹⁷ The

article also indicates that “as web archiving is a new concept to most researchers, to improve communication and promotion it is necessary to explore different metaphors that will allow different user groups to represent and imagine the web archives.”¹⁸ Still, for all that we know about what researchers might want in a web archives, what do we know about how they use what functionality already exists?

As mentioned in “Researcher Engagement with Web Archives: State of the Art,” too little is known about users’ behaviors in relation to web archives. This article was written in 2010, and since then, few articles have directly added to the literature on how researchers use web archives. This dearth of research means that

Most archiving institutions therefore rely on semi-hypothetical use cases to refine and expand their usability and interfaces. One particularly detailed study was conducted at the National Library of the Netherlands. This structured experiment, run similarly to a task-oriented usability study, evaluated user comfort level with search and access tools and attempted to determine user satisfaction with archive contents.¹⁹

The National Library of the Netherlands launched its web archiving program in 2006, the same year Archive-It debuted. The library’s collections are both domain based (collecting much of the .nl domain) and thematic in nature, and consist of collections and websites with searching that can be “carried out by URL (via the Internet Archive’s Wayback Machine) and by full-text (NutchWax).”²⁰ In 2007, the National Library of the Netherlands conducted a user survey of its web archive to form a “definitive list of user preferences, on the basis of which existing search functionality and interface can be further developed for a web archive.”²¹ During the survey, “A number of persons . . . were invited to conduct search operations in a test version of the National Library of the Netherlands web archive on the basis of concrete operations” during which it was “deliberately decided to let the test person use a full-text search option as well as a URL-based search.”²² Following the test, “users were asked about their experience of the Archive and about their preferences with regard to the selection of websites, access to the archive, and search functionality.”²³ The feedback the group received was scattered, with “a few participants” indicating that their preference would be for a full-text search option and two participants who said “they missed an advanced search function.”²⁴

In recent years, a handful of studies have looked at how historians and scholars interact with archived web pages in a variety of formats, though—to date—the field has yet to study how researchers interact specifically with Archive-It. In 2017, Susanne Belovari, assistant professor at University Archives, University of Illinois, Urbana-Champaign, carried out a thought experiment: in the year 2050, with access to only web collections from 2015, how might a historian fare if attempting

a web-based history? Belovari's imagined historian ultimately fails due to problems encountered on numerous web archiving platforms, many of which have poor usability design. While using the Wayback Machine, for instance, Belovari's imaginary user is unable to execute a keyword search, and, when using smaller collections in Archive-It, is hampered by a lack of metadata. The frustrations are fictionalized, yet speak to key questions of usability. Belovari concludes that a lack of "basic search tools and possibilities with which to scope out and identify topics, individuals, institutions, and sources"²⁵ make web research nearly impossible; furthermore, "good metadata and authority control are critical in order to find particular" web pages.²⁶ The challenges Belovari highlights do not just belong to a distant 2050 future—researchers today already attempting to use web archives encounter these same problems. Belovari's imaginary 2050 study casts light on the very real need for improvements today.

In a similar study carried out in 2007, Marilyn E. Rackley asked participants working at the North Carolina State Archives and State Library to locate specific records in either their "State Library Catalog" or the "State Government Web Site Archives Portal and the Wayback Machine."²⁷ Ultimately, Rackley's study was about context in web collections—specifically, "how easy it was to locate contextual information about the records"²⁸—but her findings, like Belovari's, shroud the task in the question of usability. In her tests, Rackley found that, often, archived websites "did not . . . contextualize [records] well enough for . . . researchers to quickly understand the record" and that "perhaps if they had more time to spend on the site as a whole, they would have demonstrated a more complete understanding" of the materials.²⁹ Providing context to help users better understand archived web pages is, of course, not a requirement of Archive-It (or of web archiving in general); however, both Belovari and Rackley illustrate various users who looked for, but could not find, information related to once-live web pages.

Though rare, usability studies focused on web archives reveal important truths about the way researchers access—or would like to access—contemporary web collections. In studying Archive-It, we aimed to address the lack of knowledge about the usability of the most widely used web archiving platform, while also contributing to the larger literature on web archives use. Our study confirms results from other web archives usability studies, contributes new findings about the importance of metadata in the usability of web archives search interfaces, and provides the first in-depth research into how users experience the Archive-It search interface in particular.

Methodology and Limitations

Our study focused on the usability of Archive-It, a tool used by approximately 94 percent of libraries and archives with web archiving programs.³⁰

Participation by repository type and user category n=14	staff	general public	library staff	students	faculty	
Governmental archives	1	1				14%
Large research university	1		3	2	1	50%
Small college			1		2	21%
Special library	1		1			14%
	21%	7%	36%	14%	21%	100%

FIGURE 5. Number and type of participants by number and type of institution

Some libraries and archival institutions use other tools for archiving the web, such as Webrecorder, but they do not provide the same level of access and discovery that Archive-It affords at this time.

Each institution represented here had a subscription to Archive-It prior to the start of this study. We conducted our testing between September and October of 2017. Our analysis does not reflect any changes made to Archive-It or the Wayback Machine since that time.

Participants

A variety of testing institutions are represented in this study, including a small college, large research universities, a governmental archives, and a special library (an art museum). We identified a range of patron profiles likely to use web archives and chose to perform user testing with representatives of those profiles. These included students, faculty members of academic institutions, members of the public, library staff members, and nonlibrarian members of institutional staff (see Figure 5 for a breakdown of participant and institution types). All of these categories were represented but not in equal proportion: library staff constituted 5 of the 14 (36%) participants, faculty 3 of 14 (21%), and nonlibrarian staff 3 of 14 (21%), with 2 students (14%) and only 1 member of the general public (7%).

Participants were recruited through email solicitations to personal networks and faculty listservs, and by word of mouth, as opposed to random sampling, because this simplified and accelerated the testing process. Each institution selected different types of users (student, faculty member, member of the public, library staff, and/or nonlibrarian institutional staff) on which to focus, and then recruited 2 or 3 users by reaching out directly to individuals via email.

These participants had varying backgrounds in web and research activity, as reported in their responses to the second question in the study (see Appendix). Based on their self-assessment, we labeled participants as “experienced researchers” or as having “limited or no archives experience.” Slightly more than half of the participants were categorized as experienced (8/14, 57%), while the remainder (6/14, 42%) were classified as having little or no archives research experience.

Future studies might benefit from different, more diverse participant pools. These could include a more balanced representation of user group segments; participants recruited in a more random way (versus through institutional and personal networks); users with a greater range of background experience, including exposure to archives generally and web archives specifically; participants from different fields of study, such as faculty in different research areas; participants from a wider public audience; and/or participants unaffiliated with the testing institution(s).

Testing Process

To ensure consistency across institutions, all researchers followed the same script (see Appendix: Script & Questions), which was based on a template from Princeton University Libraries website usability testing. Sessions were not recorded, but investigators (or their confederates) took notes on user behavior and transcribed select comments. The script had 3 sections.

BACKGROUND

We explained the purpose of the study, introduced the software, and gave a breakdown of what to expect during testing. Next, we offered assurances of anonymity and asked for an expression of consent to participate in the study. After the user consented and expressed an understanding of the study, the facilitator asked about the test participant’s use of the web and his or her archival research experience. Last, we invited first impressions of the Archive-It end-user interface before proceeding to the tasks.

TASKS

We asked participants to complete 4 exercises (find a website from a specific date, navigate through Archive-It to the Wayback Machine, identify a group of related websites, and locate a video) and then to rate their experience using the site. Our team devised these tasks based on what we consider to be core functions of Archive-It: navigating the site, successfully finding specific

captures, searching and browsing using metadata fields and facets, locating videos if captured, and interacting with the Wayback Machine. Each institution created questions that fulfilled the same aim, but tailored them to reflect its own local collections. For example, to meet the objective “A question to see if they can find a video,” Temple asked, “Can you find and watch a video from the Fellowship Farm?” while Swarthmore asked, “Can you find and watch a video from War News Radio?” The collections that we used for testing represented a variety of topics and content, and we did not require that participants have prior knowledge of our specific collections.

The 4 tasks centered on finding specific resources (e.g., locate the City Parks Association website in June 2013) rather than general browsing behavior. The exercises were designed to be straightforward tasks with clear outcomes (completed or not completed), resulting in a small and quantifiable pool of data to analyze. More in-depth questions and forms of analysis could be explored in an expanded study.

CONCLUSION

Last, we asked users to rate the site overall on a scale of 1 to 10. They were also invited to provide concluding thoughts and offer any suggestions for improving the Archive-It interface, and offered a chance to ask their own questions.

Analysis Process

Because we sought a tool that would allow easy exchange between our 7 collaborators at geographically distinct locations with different capacities to purchase software, we decided to do all of our analysis using Google Sheets.

Initially, we planned to analyze quantitative results of each task based on whether or not the participant completed the task. However, we did find some gray areas while testing and developed 3 categories for results: completed easily, completed with difficulty, or did not complete. “Completed easily” indicates the user was able to navigate the site and successfully complete the task without any assistance. “Completed with difficulty” means that the user did eventually complete the task, but encountered significant issues and confusion, and perhaps received a minor hint from the test administrator. “Did not complete” indicates the user could not complete the task within approximately 10 minutes, or became frustrated and gave up. We did not collect data on task completion time, but estimate that based on the 30 to 45 minutes we told participants the testing would take, if a participant could not complete the task in less than

10 minutes, the test facilitator moved on to the next task and considered the previous task a “Did not complete.”

Through an iterative and collaborative process of qualitative analysis, we created 24 codes recognizing the most common behaviors or themes, such as sorting search results successfully or unsuccessfully, using the search bar appropriately or incorrectly, and understanding versus being confused by faceted filtering of results. Each test administrator coded her own notes from the testing she conducted at her institution. We created charts and graphs to analyze and visualize certain types of data. We also compiled suggestions for the Internet Archive and for archivists using the Archive-It service, based on users’ direct feedback or our observations based on their behavior.

Concerns

We recognize that bias is inherent in every research design, and we tried to be aware of bias during this study.

Our test instrument reflects a particular understanding of archives and digital libraries, influenced by the papers referenced in our literature review and characterized by certain shared ideas: the concept of archival collections and how they are organized via Archive-It, familiarity with metadata and description practices in archives and libraries, and the idea of websites as archival artifacts. Investigators with different backgrounds who came to the study without these same preconceptions about archives and digital libraries may have formulated aspects of the study differently.

In our analysis, we did not account for the computer or information literacy of participants, variation in user profile (faculty, students, public, or staff), or variation in the type of institution (academic, state archives, or museums). Accounting or controlling for these variables might have changed our analysis.

Findings: Task Descriptions and Results

The following section reports task completion rates and participants’ interface ratings and analyzes these results as well as the frequency of the codes investigators applied to the data.

Overall, participants performed well and completed most tasks, as shown in Figure 6. Nearly half of the participants (6 of 14) completed all of the tasks. An equal number completed half of the tasks. One participant completed 3 of the 4 tasks. Only 1 participant was unable to complete 3 tasks, and no participants were unable to complete all 4 tasks.

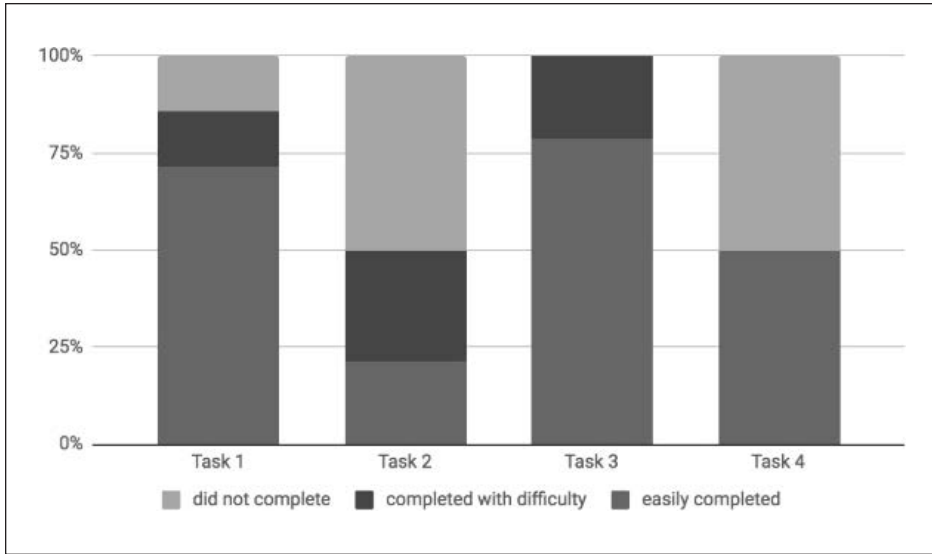


FIGURE 6. Task completion rate

TASK 1: FIND A CAPTURE IN ARCHIVE-IT

In the first task, test facilitators asked participants to find a web page captured on a certain date in Archive-It (e.g., “Can you find the Swarthmore College homepage in March 2011?”). Testers used a variety of methods to find the websites they were asked to locate. Figure 7 shows some of the strategies employed: the Archive-It collection search bar at the top of the screen, the facet filters on the left panel, and browsing seed metadata toward the bottom right.

Of the 14 testers, 12 successfully located the exact version of the site that they were asked to find (85%). Most used the Archive-It collection search bar to locate a seed; this approach was successful 6 times out of 8. Testers attempted to use facets to identify the desired seed and date 6 times, but were only successful twice. In most cases, using faceted filters did not lead to the results the testers wanted or expected: either the facets did not narrow the results enough, or they eliminated the desired result because the participant did not realize a secondary filter was still applied, or the participant did not apply the archivist-provided metadata as expected (for example, dates added by an archivist in the “date” field did not necessarily reflect the capture dates). Six participants looked at the dates in the metadata entries for a website in Archive-It (either the descriptive metadata entered by an archivist, or the “Captured” dates automatically generated by Archive-It and added to the metadata section) to find the range for the desired date. These participants had a higher success rate, leading to a desired outcome half of the time. Other



FIGURE 7. Collection landing page

strategies included leaving the Archive-It site to utilize the Internet Archive’s Wayback Machine. Three participants successfully used this strategy 4 times; 1 participant was unsuccessful in trying this strategy. The least-used strategy was searching via methods other than the Archive-It search engine. Two participants were successful with CTRL+F, and one was unsuccessful with a Google search. The results of this task illustrate that participants relied on a variety of search strategies in assorted combinations.

TASK 2: FIND A CAPTURE IN THE WAYBACK MACHINE

The second task asked participants to find captures from dates that were not available through the repositories’ Archive-It captures for that website, but were available in the general Internet Archive through the Wayback Machine



FIGURE 8. Wayback Machine homepage



FIGURE 9. Seed metadata including link to finding aid in Description

(i.e., <https://archive.org/web>, Figure 8). This task tested whether people were aware of the Wayback Machine and how to access it. This was the most difficult task for participants in all groups. Half of them failed, while 4 completed it with difficulty. Only 3 participants easily completed the task.

Using the search bar for an Archive-It collection was the most commonly used strategy, and it was unsuccessful most of the time. Only 4 times did it result in participants finding information that assisted them in navigating to the website they were asked to find. The second most-used strategy was to go directly to the Wayback Machine. Used 8 times, it was only successful for participants in 3 cases. Other strategies were used much less frequently, although 4 participants employed Google search and CTRL+F with some success in getting to a desired result or to the Wayback Machine.

One participant, familiar with finding aids, followed a link in an Archive-It seed's descriptive metadata to an associated finding aid (see Figure 9). However, the finding aid did not include a link to the Wayback Machine, so, ultimately, referring to the finding aid did not assist the participant in completing this task.

Four participants tried strategies such as faceting (3 times) and sorting (1 time), but these methods, although used correctly, were not helpful in completing the task. The low task completion rate illustrates that many participants may not have been aware of web archive discovery resources (such as the Wayback Machine) beyond what they learned from participating in this study.

TASK 3: COUNT RELATED SEEDS

In the third task, participants had to identify a group of websites related by creator, theme, or some other commonality, such as “How many websites are based in Swarthmore, PA?” This appeared to be the easiest task for all participants to complete. This task was easily completed by 11 participants, while 3 completed it only with difficulty. No one was unable to complete this task. The most popular approach to answering this question was to use the Archive-It search; it was employed 15 times, with 7 of those being successful. Of the 8 unsuccessful uses of search, most involved the participant either not understanding the terminology employed by Archive-It, or, in the case of 1 participant, not realizing that he or she was searching the full text of the websites instead of just the Archive-It metadata fields. Other successful approaches included clicking on the facets (7 times) and looking at the descriptive metadata (5 times). Facet was successful 5 times and metadata 4 times. The results of this task show that Archive-It does a good job of relating grouped materials, and most participants were successful at finding materials related by creator or theme.

TASK 4: FIND A VIDEO

The last task asked participants to find a video within an Archive-It collection. Half of the participants completed it easily, while the other half failed completely. Following the link in seed metadata, as shown in Figure 10, was the most popular strategy (21 times). Participants also used the main search bar (15 times), navigating through Wayback (1), a Ctrl+F search on the collection page, and facets (3). Four participants were able to locate where videos would have been on the live site by browsing the archived site, but discovered that Archive-It had not captured them. Those who used metadata to find videos were successful 15 times out of 21. Searching was successful 11 out of 15 times.

Some of the difficulty with this task was due to the videos not being successfully captured by Archive-It, although the participants otherwise succeeded in locating where the videos would have been on the live site. Results of this task show that Archive-It does not do an adequate job of situating videos within its website context, and a significant number of participants, possibly 50%, missed

Page 1 of 1 (1 Total Results)

Sort By: Best Match | Title (A-Z) | Title (Z-A) | URL (A-Z) | URL (Z-A)

Title: The Princeton Triangle Club Public Website

URL: <http://www.triangleshow.com/>

Description: The Princeton Triangle Club Public Website provides information on the activities of the university's oldest theatrical club, including a list of trustees, officers, and staff, and video clips of productions showing song and dance, sketches, kicklines, reunions, behind-the-scenes, vintage footage, and full-length videos of select shows. The site also includes information about ticket sales, merchandise, how to join the club, how to become a supporter, and links to social media. To find more archival records of the Princeton Triangle Club, visit the collection's finding aid at <http://findingaids.princeton.edu/collections/AC122>.

Captured 46 times between Jul 5, 2016 and Aug 22, 2018

Videos: 8 Videos Captured

Group: Student Life

Creator: [The Princeton Triangle Club](#)

Collection Call Number: AC122

Page 1 of 1 (1 Total Results)

FIGURE 10. Seed metadata including link to videos

SitesVideosSearch Page Text

The following unique videos were discovered from <http://warnewsradio.org/>

Page 1 of 3 (242 Total Results) [Next Page ►](#)

Sort By: Best Match

Video URL: <https://rs--sn-a5mekney.googlevideo.com/videoplayback?expire=1476857440&ra...mn=sn-a5mekney&ms=ltu&mt=1476835722&mv=m&nh=IgpwcjAyLnNqYzA3KglxMjcuMC4wLjE>

Archived From URL: <https://www.youtube.com/watch?v=pZT2BpH1Lrs>

Watch

Videos Captured on Oct 19, 2016

Type: video/webm

Size: 32 MB

Video URL: <https://r6--sn-a5mekn7d.googlevideo.com/videoplayback?mime=video%2Fwebm&ke...mn=sn-a5mekn7d&ms=ltu&mt=1476847285&mv=m&nh=IgpwcjAyLnNqYzA3KglxMjcuMC4wLjE>

Archived From URL: <https://www.youtube.com/watch?v=oKh6Kzqc428>

Watch

Videos Captured on Oct 19, 2016

Type: video/webm

Size: 15 MB

FIGURE 11. Video metadata view in Archive-It

otherwise viewable video content. Several participants specifically complained that the metadata for the videos was confusing or inadequate.

Following the seed metadata was the most popular strategy for finding videos. This is in contrast to tasks 2 and 3, where search was the most popular strategy, or task 1, where strategies were evenly distributed. In this particular case, this may have been because the metadata for each seed described the content requested by the task, limiting the need for extensive use of the search function.

Participants’ Interface Ratings

At the conclusion of testing, participants were asked to rate their overall impression of the Archive-It end-user interface in terms of its user-friendliness on a scale of 1 to 10 (see Figure 12). The average rating was a 6.5, with 1 participant giving a rating of 3, 1 participant giving a rating of 9, and the rest of the participants clustered in the 5 to 8 range. The mode (most frequently given) score was a 7.

The outlier low score of 3 may indicate that this participant especially struggled with the tasks. While all of the other participants in our study were able to complete between 2 and 4 tasks, the participant who gave a rating of 3 was the only one unable to complete 3 out of 4 tasks.

Analysis

In addition to assessing each task, the research team considered the overall strategies that participants used when attempting to complete each task. This

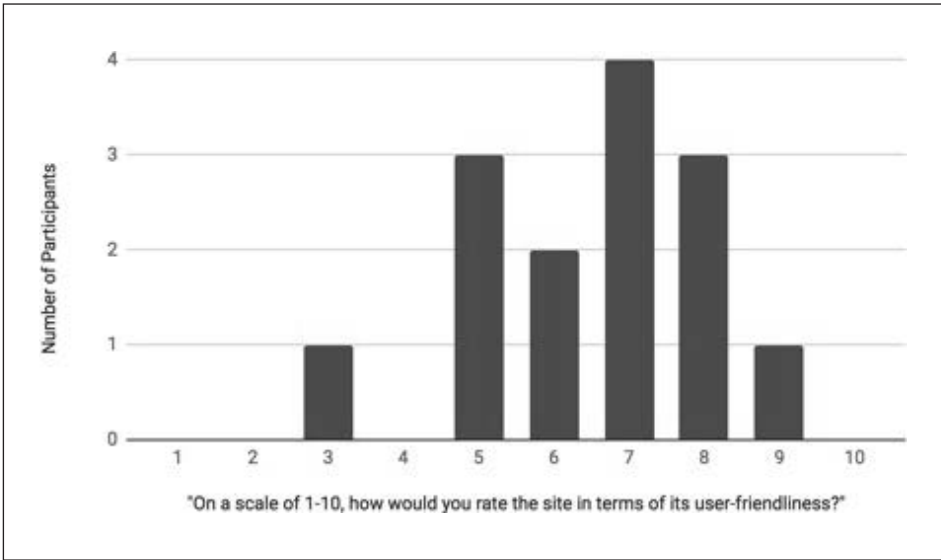


FIGURE 12. Overall rating of interface

analysis provides an overview of the strategies that were generally successful or unsuccessful, independent of each task.

Nine of the participants were confused when using the search bar. Unsure of the scope and precision of the search, they nonetheless expressed expectations for an intuitive and comprehensive search. Without taking the time to formulate a search strategy, participants tried entering many different types of information into the search bar: numerical dates and even analytic terms like “main site” or “homepage” in addition to keywords. A comprehensive search algorithm integrating both metadata and full-text search with appropriately weighted relevancy, or more obvious differentiation between the 2, would benefit these participants. However, this criticism is softened by the observation that successful use of the search bar was tagged nearly as often: 31 times. So, while the search bar caused confusion for some participants, in nearly as many occurrences, participants used it successfully. Making the change suggested above, if done properly, should not negatively affect participants who used the search bar successfully.

We also coded the following in our analysis: misleading or unsatisfactory metadata display, confusion regarding the meaning of collection (a group of websites) versus seed (a single website), and misuse of facets or expressing the view that the facets were unhelpful or confusing. It is worth noting that for the number of times each attribute or tool was unsuccessful in helping a participant complete a task, it sometimes helped different participants, or even the same participant in a different moment. Metadata (reading, understanding, and clicking on the descriptive metadata displayed with each seed on the collections page) was helpful 3 times as often (38) as it was unhelpful or confusing. Facets (using the faceted terms along the left side of the collections page as filters) was helpful about the same number of times (14) as it was unhelpful or confusing. All references to collections in the data caused confusion. Participants expressed confusion as to the meaning of an Archive-It collection versus a repository’s holdings or a traditional analog collection—demonstrating that the organizational structure of collections in Archive-It tends to make users feel more, rather than less, disoriented.

Participants raised half of the metadata complaints on the video task, including multiple complaints about the lack of information, or even a title, for each video, and confusion as to which link to click on in results. Several participants also expressed confusion over the meaning of some metadata terms—“creator” versus “publisher,” and the meaning of “coverage.” Some participants did not notice metadata in Archive-It that would have led them directly to videos. In some cases, facets caused frustration for participants who struggled with the search because they did not realize a filter was still applied, and their results were unknowingly restricted.

A common point of confusion was the terminology used by Archive-It to label the organization of the websites. Nine participants explicitly voiced confusion over the meaning of the terms “capture,” “coverage,” and “creator,” as well as “collection” versus “site” or “archival collection,” and “title” versus “URL.” That some of these terms come directly from Dublin Core standards does not change the fact that the terms confused many participants. It would be more helpful to users if archivists hid or changed the labels, or defined certain metadata fields.

Five participants were confused about the concept of a “collection” in Archive-It. Four also expressed the view that the arrangement of the site was confusing, and had difficulty navigating from a seed page to a collection page to an organization page and understanding the differences between these entry points.

In 10 occurrences, participants used an alternative search, such as Google or the CTRL+F in-browser search function. Six times this strategy was unsuccessful, and the other four times, the users succeeded using alternative searches.

On 6 occasions, participants encountered difficulty navigating between the subscription Archive-It archive and the general Internet Archive collection. Nearly as many times, a total of 4, participants were successful. In only 1 instance was the participant unable to navigate the dates in the Wayback Machine, where 4 attempts were successful.

Only Swarthmore’s 3 participants encountered at least a moment when they accidentally navigated to the live site without realizing they had left the archive. This may have been because the Swarthmore tasks used the college’s homepage, a website all test participants were familiar with; other institutions’ tasks used web pages that test participants likely had rarely or never seen before.

Participants made general positive remarks about site design and function. One participant liked the relaxing color scheme and felt the site was welcoming and clean; a second echoed that the design was clean and well laid out. Several remarked that it could be a useful reference tool or that they never thought about websites being archived but were pleased to know that they were.

There were also general negative remarks. As many participants commented that the general design was too busy, too filled, or overwhelming, as had commented that the general design was clean. Participants remarked that the site was disorganized; one wished the search bar was higher and another wished the research content and collections came before any information about the partner organization. Another participant felt that too many click-throughs were required to get to the pertinent information.

A recurring theme of this user study is that Archive-It is initially unintuitive, but that once a user becomes acclimated to its layout, logic, and terminology, it becomes easy to use. Eight of the participants in the study expressed this sentiment, which we characterized as “learning curve.” One person stated

that a librarian may find it easy to navigate the site, but that a general user of the Internet may struggle. Another felt he or she would have to try certain things first to figure it out. However, as one participant commented, “After a day of using it, I wouldn’t have trouble.” Someone else noted, “once you get into it, it works well.” A different user similarly stated, “once [he or she] understood how the site worked it was easy to navigate.”

Recommendations

The resulting data set of this study identified 6 key areas for improvement to the Archive-It interface: the search bar, video access, terminology glossary, date facet, published metadata options, and metadata labels. Our recommendations for clarifying these issues follow.

DEVELOP A MORE COMPREHENSIVE ONE-SEARCH BOX

The attribute that most inspired frustration was the search bar. There, our primary recommendation for improvement would be to develop a clearly comprehensive one-search box, integrating both facets and full-text search. Short of changing Archive-It’s search algorithm, an explanation or clearer indication of what exactly is being searched within the search box (such as added metadata versus website text) and what is not being searched (such as when a filter is applied) could help users to more effectively search within the interface.

Given that a portion of the participants in this study successfully used the search bar for task completion, it is imperative that making the changes suggested not negatively affect such users.

ADD METADATA FOR VIDEOS

Another clear takeaway is that researchers could benefit from the addition of metadata for videos, such as the video title or at least a filename.

ADD TOOLTIPS OR A GLOSSARY TO EXPLAIN POTENTIALLY UNFAMILIAR TERMS

Many users found terminology within the interface to be confusing (such as the term “coverage,” which usually means geographic area; or the relationship between a URL and a collection). That some of these terms come directly from Dublin Core standards does not change the fact that the terms confused many researchers. Integrated pop-ups/tooltips with definitions of potentially unfamiliar terms—specifically written in plain language free of library/archives

or internet jargon—could be a great improvement. Additionally, incorporating tooltips to explain the best use of the search functionality within the site could mitigate confusion and set expectations for users as to how the search bar is designed. If tooltips are not possible, providing an easy-to-find glossary of terms on the end-user interface might provide the answers some users need.

NORMALIZE AND INDEX DATE FIELD

The study data also suggest that the system's current configuration to index entries in the "date" field as text strings is confusing and unhelpful to researchers hoping to use the date facets. Instead, common date formats should be recognized, normalized, and parsed for effective scoping or faceting. Optimally, Archive-It's date field should offer both an ISO 8601-compliant date entry in W3CDTF format (YYYY-MM-DD), to be indexed, as well as a nonindexed date expression field.

UNPUBLISH METADATA OPTION

Our user study showed that some of the displayed metadata fields were confusing, or at least unhelpful, and not used by most researchers. However, archivists may wish to use these fields for internal purposes and recordkeeping. Currently, the ability to suppress metadata fields exists, but the functionality is only available to administrators and 5 of the 7 archivists conducting this study were not aware of the option. Giving archivists the ability to more easily set particular metadata fields to "unpublished" via a simple toggle tool on the staff interface for metadata entry could be advantageous to both archivists and end-users.

MODIFYING METADATA LABELS

Currently, archivists can add new metadata fields, but they do not have the ability to change how a metadata field is labeled (for example, one might prefer to label the "Coverage" field "Geographic Coverage" instead). Customized metadata labels could make the contents of certain fields clearer to researchers.

Conclusion

Comprehensive and intuitive web archiving tools are essential to support the ongoing work of archivists who increasingly rely on this technology to preserve and access websites. Without them, important historical information

would be unavailable after websites are updated, edited, or removed from the Internet. More robust data about the usability of web archiving will support the development of tools that meet the needs of both archivists and researchers.

In our study, participants' overall impressions of Archive-It were positive-leaning, with an average score of 6.5 out of 10. Archive-It was not initially intuitive for most participants, though many noted growing ease as they overcame the learning curve and became acclimated to the site. The study found that individuals who self-reported as being experienced in doing archival research were more successful in completing the tasks overall compared to participants lacking experience with researching in archives. We discovered that some of the terminology employed by Archive-It confused participants, which we suggest could be mitigated with a glossary, tooltip definitions, or the ability to customize metadata labels. We also noticed that videos were difficult for participants to find and access through the Archive-It interface due to unsuccessful capture or insufficient metadata. Half of all metadata complaints were associated with the video task; the addition of video metadata could improve users' experience. In addition to analysis of Archive-It, the data collected in this study offer insight into the behaviors and expectations of internet users. For example, we found significant search bar confusion as participants expected a comprehensive single-search algorithm. They also expected date expressions, currently indexed as strings in Archive-It, to be normalized, indexed as dates, and searchable.

While the data compiled in this study were instructive, future studies using larger, more diverse pools of participants and institutions are needed to fully explore the usability of Archive-It and other web archiving tools.

Appendix: Tasks Script and Questions

(Authors' Note: To view study results summary, codebook, notes from testing sessions, and bibliography, please visit <http://arks.princeton.edu/ark:/88435/dsp01d791sj97j>.)

Introduction

Thank you so much for coming in today to help test the website of Archive-It, our web archiving subscription service. I wanted to give you a little information about what you will be looking at and give you time to ask any questions you might have before we get started.

Test Participant's Role

As mentioned, we use a software service called Archive-It from the Internet Archive to archive websites. I am going to ask you to complete a set of tasks that people often use web archives to accomplish. Our goal is to see how easy or difficult you find the site to use. Nothing is designed to be a trick question, and there are no right or wrong answers. We're not trying to judge your abilities in using these services; rather, we're interested in seeing how well the website works for people who come to the site for a variety of reasons and with different backgrounds and different levels of experience using archived websites.

As you use the site, please do so as you would at home or your office. During the session, I would like you to think aloud as you work to complete the tasks. Try to keep a running dialogue of what you're thinking and seeing, why you clicked on something or chose a particular approach to an assigned task, and any reactions you have to the site (for example, surprise, confusion, delight, annoyance, etc.). Any feedback you can give us will be very helpful as we work to improve the site, so please be open and direct in sharing your thoughts, whether positive or negative.

As you work through the tasks, try to stay within the Archive-It website, but if you reach a point where you are not sure where or how to find something, please feel free to use any method you'd normally use, including Google.

Test Facilitator's Role

I will not offer any suggestions or hints on the tasks, but if you ever feel that you don't understand what you're supposed to accomplish in a task, please let me know, and I'll try to clarify. From time to time, I may ask you some questions to dig deeper into why you clicked or searched in a particular way or to clarify your expectations or reactions.

A colleague will be observing and taking notes while you complete the tasks. Your name will not be associated or reported with data or findings from this evaluation. As I mentioned, we are not testing you, we are observing you

as you test the site. Are you okay with us taking notes and using the data we collect from this session in a public white paper?

This session should last between 30–45 minutes.

Do you have any questions for me before we start?

Background

I'd like to start by learning a little bit about you and your work.

1. Can you tell me a little bit about your use of the Web. Which websites do you use on a regular basis, and how would you characterize the amount of time you spend online doing work?
2. Do you have any experience doing archival research? If so, what was that like? Have you ever looked at or used archived websites for research before? If so, what was your experience like?
3. First impressions of the Archive-It site: Take a minute to look at this page and tell me what impression you get from it: what you think a person might be able to do on this site, and if anything strikes you about it. Just explore the page and tell me what you're noticing even if it seems like it would be obvious to anybody.

Tasks/Questions

I'd now like to ask some questions about finding material on the site:

1. A question to make them check a specific date:
 - a. Swat: Can you find the Swarthmore College homepage in March 2011?
 - b. Temple: Can you find the City Parks Association website in June 2013?
 - c. Princeton: Can you find the Princeton University History department site from 2016?
 - d. Delaware: Can you find (former) Governor Markell's homepage from August 2012?
 - e. Frick: Can you find the Museum of Modern Art website from May 2015?
 - f. Ivy Plus: Can you find Evan Ziporyn's website from October 2015?
2. A question to see if they realize there are other dates in general Internet Archive not in our Archive-It:
 - a. Swat: What did the Swarthmore homepage look like in 1997?
 - b. Temple: What did the AIDS Library website look like in 2000?
 - c. Princeton: What did the Princeton homepage look like in 2000?
 - d. Delaware: What did the Town of Magnolia's website look like in 2003?
 - e. Frick: What did the Brooklyn Museum homepage look like in 2002?
 - f. Ivy Plus: What did Evan Ziporyn's website look like in 2006?
3. A question to see if they can identify a group of related websites (same creator, participant theme, etc.):

- a. Swat: How many websites are based in Swarthmore, PA?
 - b. Temple: Can you find the websites related to Mount Airy?
 - c. Princeton: Can you locate all of the Political Science department's archived websites?
 - d. Delaware: Can you locate websites related to DelDOT?
 - e. Frick: Can you locate websites of New York City art galleries?
 - f. Ivy Plus: Can you find websites related to Hurricane Sandy?
4. A question to see if they can find a video:
 - a. Swat: Can you find and watch a video from War News Radio?
 - b. Temple: Can you find and watch a video from the Fellowship Farm?
 - c. Delaware: Can you find and watch a video from Delaware State Courts?
 - d. Princeton: Can you find and watch a video of a Triangle Club performance?
 - e. Frick: Can you find and watch a video about the Lenox Pediment Clock?
 - f. Ivy Plus: Can you find and watch a video from Thea Musgrave?
 5. Do you have any other comments or questions?

Conclusion

That's the end of the tasks. Based on your experience in this session, I want to get your overall impression of the site. On a scale of 1–10, how would you rate the site in terms of its user-friendliness? Are there any changes you would recommend to improve the site?

Before you go, do you have any comments or questions about the site?

Thanks for coming in today!

NOTES

¹ Matthew Farrell et al., "Web Archiving in the United States: A 2017 Survey," National Digital Stewardship Alliance (October 2018), 20, <https://osf.io/ht6ay>.

² Archive-It, "About Us," last modified 2014, <https://archive-it.org/learn-more>.

³ Jefferson Bailey et al., "Web Archiving in the United States: A 2016 Survey," 4, https://ndsa.org/documents/WebArchivingintheUnitedStates_A2016Survey.pdf.

⁴ Bailey et al., "Web Archiving in the United States," 4.

⁵ Ferrell et al., "Web Archiving in the United States: A 2017 Survey," 20.

⁶ Archive-It, "About Us."

⁷ *Archive-It User Guide*, "Add, Edit, and Manage Your Metadata," last modified September 2018, <https://support.archive-it.org/hc/en-us/articles/208332603-Add-edit-and-manage-your-metadata>.

⁸ Shawn M. Jones, Michael L. Nelson, Alexander Nwala, and Michele C. Weigle, "The Many Shapes of Archive-It," Cornell University (June 2018): 3, <https://arxiv.org/abs/1806.06878>.

⁹ Jones et al., "The Many Shapes of Archive-It," 1.

¹⁰ Rachel Walton, "Looking for Answers: A Usability Study of Online Finding Aid Navigation," *American Archivist* 80, no. 1 (2017): 30–52, <https://doi.org/10.17723/0360-9081.80.1.30>; Elizabeth Yakel, "Encoded Archival Description: Are Finding Aids Boundary Spanners or Barriers for Users?," *Journal of Archival Organization* 2, nos. 1–2 (2004): 63–77, https://doi.org/10.1300/J201v02n01_06; Xiaomu

- Zhou, "Examining Search Functions of EAD Finding Aids Web Sites," *Journal of Archival Organization* 4, nos. 3–4 (2007): 99–118, https://doi.org/10.1300/J201v04n03_06; Burt Altman and John R. Nemmers, "The Usability of On-Line Archival Resources: The Polaris Project Finding Aid," *American Archivist* 64, no. 1 (2001): 121–31, <https://doi.org/10.17723/aarc.64.1.80300272655rqu74>; Christopher J. Prom, "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 67, no. 2 (2004): 234–68, <https://doi.org/10.17723/aarc.67.2.7317671548328620>; Morgan G. Daniels and Elizabeth Yakel, "Seek and You May Find: Successful Search in Online Finding Aid Systems," *American Archivist* 73, no. 2 (2010): 535–68, <https://doi.org/10.17723/aarc.73.2.p578900680650357>; Lora J. Davis, "Providing Virtual Services to All: A Mixed-Method Analysis of the Website of Philadelphia Area Consortium of Special Collections Libraries (PACSCL) Member Repositories," *American Archivist* 75, no. 1 (2012): 35–55, <https://doi.org/10.17723/aarc.75.1.a716w067468262h5>; Magia Ghetu Krause and Elizabeth Yakel, "Interaction in Virtual Archives: The Polar Bear Expedition Digital Collections Next Generation Finding Aid," *American Archivist* 70, no. 2 (2007): 282–314, <https://doi.org/10.17723/aarc.70.2.lpq61247881t10kv>; Zhou, "Examining Search Functions of EAD Finding Aids Web Sites."
- ¹¹ Helen Hockx-Yu, "Access and Scholarly Use of Web Archives," *Alexandria* 25, nos. 1–2 (2014): 113, <https://doi.org/10.7227%2FALX.0023>.
 - ¹² David Cruz and Daniel Gomes, "Adapting Search User Interfaces to Web Archives," in *Proceedings of the 10th International Conference on Preservation of Digital Objects*, ed. José Borbinha, Michael Nelson, and Steve Knight (Lisbon: National Library of Portugal, 2013), 326–9, <https://sobre.arquivo.pt/wp-content/uploads/adapting-search-user-interfaces-to-web-archives.pdf>.
 - ¹³ Arthur Thomas et al., "Researcher Engagement with Web Archives: Challenges and Opportunities for Investment," *Joint Information Systems Committee Report* (August 2010): 8, <https://ssrn.com/abstract=1715000>.
 - ¹⁴ Cruz and Gomes, "Adapting Search User Interfaces to Web Archives," 1–2.
 - ¹⁵ Cruz and Gomes, "Adapting Search User Interfaces to Web Archives," 2.
 - ¹⁶ Cruz and Gomes, "Adapting Search User Interfaces to Web Archives," 4.
 - ¹⁷ Peter Stirling, Philippe Chevallier, and Gildas Illien, "Web Archives for Researchers: Representations, Expectations and Potential Uses," *D-Lib Magazine* 18, nos. 3–4 (2012), <http://www.dlib.org/dlib/march12/stirling/03stirling.html>.
 - ¹⁸ Stirling et al., "Web Archives for Researchers."
 - ¹⁹ Meghan Dougherty et al., "Researcher Engagement with Web Archives: State of the Art," *Joint Information Systems Committee Report* (August 2010): 24, <https://ssrn.com/abstract=1714997>.
 - ²⁰ Marcel Ras and Sara van Bussel, "Web Archiving User Survey," National Library of the Netherlands (July 2007), 7, https://www.kb.nl/sites/default/files/docs/kb_usersurvey_webarchive_en.pdf.
 - ²¹ Ras and van Bussel, "Web Archiving User Survey," 7.
 - ²² Ras and van Bussel, "Web Archiving User Survey," 13.
 - ²³ Ras and van Bussel, "Web Archiving User Survey," 13.
 - ²⁴ Ras and van Bussel, "Web Archiving User Survey," 13.
 - ²⁵ Susanne Belovari, "Historians and Web Archives," *Archivaria* (Spring 2017): 68, <https://archivaria.ca/index.php/archivaria/article/view/13600>.
 - ²⁶ Belovari, "Historians and Web Archives," 72.
 - ²⁷ Marilyn E. Rackley, "Navigating the Web Archives: A Study of Users' Understanding of Context" (master's thesis, University of North Carolina, July 2007): 41, <https://cdr.lib.unc.edu/record/uuid:babf9e65-7bb1-4f28-81da-8cb95272c823>.
 - ²⁸ Rackley, "Navigating the Web Archives," 59.
 - ²⁹ Rackley, "Navigating the Web Archives," 59.
 - ³⁰ Farrell, "Web Archiving in the United States: A 2017 Survey," 20.

ABOUT THE AUTHORS



Samantha Abrams is the web resources collection librarian for the Ivy Plus Libraries Confederation, where she works with librarians to create accessible, thematic collections of at-risk web content. She holds a master's degree in library and information studies from the University of Wisconsin–Madison, and a bachelor's degree in English and creative writing from the University of Iowa.



Alexis Antracoli is assistant university archivist for technical services at Princeton University Library where she provides leadership for the Archival Description and Processing Team. She also teaches at the Rutgers University School of Communication and Information Studies. Her professional interests include applying user-experience research and user-centered design to archival discovery systems, developing and applying inclusive description practices, and web archiving. She holds an MSI in archives and records management from the University of Michigan, a PhD in American history from Brandeis University, and a BA in history from Boston College.



Rachel Appel is the digital projects and services librarian at Temple University. She manages and oversees digital collections and metadata projects related to cultural heritage and scholarly communications. She manages the DPLA hub PA Digital, the implementation of a research information management system across campus, as well as several other digital projects. Appel received her MSIS from the University of Texas at Austin with a specialization in digital archives and preservation and her BA from Smith College in film studies.



Celia Caust-Ellenbogen is an archivist at the Friends Historical Library of Swarthmore College.



Sarah Denison is Delaware's deputy state archivist and manages the state's archives, records management, and historical markers programs. She earned her MLIS from Drexel University, BAs in history and history of art and architecture from the University of Pittsburgh, and is a member of the Academy of Certified Archivists. Before joining the archival profession, Denison worked in the newspaper industry and as a private investigator. These experiences influence her professional interests, which include expanding public access to data and increasing public engagement and transparency throughout government.



Sumitra Duncan is head of the web archiving program at the Frick Art Reference Library. In this role she manages the web archiving program of the New York Art Resources Consortium (NYARC), which consists of the Frick Art Reference Library and the libraries and archives of the Brooklyn Museum and the Museum of Modern Art. She cofounded and cocoordinates the Web Archiving Special Interest Group (SIG) of the Art Libraries Society of North America (ARLIS/NA) and the Archive-It New York Users Group. She holds an MSLIS from Pratt Institute and a BA in English from Virginia Commonwealth University.



Stefanie Ramsay is the digital projects librarian at Temple University. She oversees a portfolio of digitization and metadata creation projects in coordination with Temple's archival repositories and serves as the co-project manager for the DPLA hub for Pennsylvania, PA Digital. Ramsay received her MLIS from the University of Washington and her BA in history from Occidental College.