Routledge
Taylor & Francis Group

Journal of Web Librarianship

ISSN: 1932-2909 (Print) 1932-2917 (Online) Journal homepage: https://www.tandfonline.com/loi/wjwl20

Accessing Web Archives: Integrating an Archive-It Collection into EBSCO Discovery Service

Christina A. Beis, Kayla Nicole Harris & Stephanie L. Shreffler

To cite this article: Christina A. Beis, Kayla Nicole Harris & Stephanie L. Shreffler (2019) Accessing Web Archives: Integrating an Archive-It Collection into EBSCO Discovery Service, Journal of Web Librarianship, 13:3, 246-259, DOI: 10.1080/19322909.2019.1625844

To link to this article: https://doi.org/10.1080/19322909.2019.1625844

JOURNAL OF WEB LIBRARIANSHIP
2019, VOL. 13, NO. 3, 246–259
https://doi.org/10.1080/19322909.2019.1625844

Routledge
Taylor & Francis Group

Accessing Web Archives: Integrating an Archive-It Collection into EBSCO Discovery Service

Christina A. Beis[a], Kayla Nicole Harris[b], and Stephanie L. Shreffler[a]

[a]Roesch Library, University of Dayton, Dayton, USA; [b]Marian Library, University of Dayton, Dayton, USA

ABSTRACT

Effective collaboration between archives and technical services can increase the discoverability of special collection materials. Archivists at the University of Dayton Libraries began using Archive-It to capture websites relevant to their collecting policies in 2015. However, the collections were only made available to users from the University of Dayton page on the Archive-It website. Content was isolated in a separate platform and was not promoted to users. Working together, the team of archivists and technical services librarians incorporated the web archive collections into the Libraries' EBSCO Discovery Service (EDS) discovery layer. A local data dictionary was created based on OCLC's *Descriptive Metadata for Web Archiving* report (2018), and metadata was added at the seed and collection levels. The result was indexed content on a single, user-friendly platform. The web archive collections were then marketed to the University of Dayton community, and statistics were generated on their use.

KEYWORDS

web archiving; discovery layer; special collections; collaboration; social media; academic libraries; EBSCO Discovery Services; web-scale discovery; Archive-It; metadata

Introduction and background

Web archiving is an emerging area of practice within libraries and archives that involves the process of selecting, capturing, and making accessible content from the web. Archivists and special collections librarians are responsible for preserving material of enduring value, so naturally this should extend to digital resources, including information that is only available online. Websites, similar to other born-digital content, provide a unique set of challenges but web archiving follows much of the workflow for traditional archival work. Just like traditional archival materials, websites need to be collected according to a collection development policy, arranged, and described to provide context for the user, and made available for use.

The Internet Archive, a nonprofit organization based out of California, has been capturing websites since 1996 as part of its mission to provide access to all knowledge. As the pioneer in the field of web archiving, its automated crawls collect around 1.5 billion pages per week and about 1 million captures

CONTACT Christina A. Beis cbeis1@udayton.edu Roesch Library, University of Dayton, 300 College Park, Dayton, 45469, USA.

per week from individuals using the "Save Page Now" function of its "Wayback Machine" (Matisse, 2018). This tool allows anyone to capture an individual website page and to add it permanently to the Internet Archive's collection of websites. In addition to being used by cultural heritage institutions, many Internet users are familiar with the Wayback Machine and it is frequently referenced in popular media as a trusted source for referring back to previous iterations of websites or deleted content.

In 2006, the Internet Archive launched Archive-It, a subscription-based web archiving service for cultural heritage institutions. Institutions access Archive-It through a web application to harvest, organize, and catalog their collection material. Users can automate the crawls with different frequencies, set parameters about the depth that the crawls will take, and add metadata at the item and collection levels to enable better searching for users. All content is stored at Internet Archive data centers, and the annual subscription is based on the amount of storage space needed. Each subscribing institution's collections are available through a page on the Archive-It website (https://archive-it.org) that features a search box and facets for users to navigate, as seen in Figure 1.

The University of Dayton Libraries purchased a subscription to Archive-It in 2014. By early 2015, archivists from the Libraries' three special collections units began capturing content relevant to each collection. The University Archives and Special Collections contains records of University departments, student organizations, faculty papers, and other special collection material. The U.S. Catholic Special Collection preserves records of the Catholic Church and Catholic life in the United States, while the Marian Library collects material that documents the diverse expressions of popular devotion to Mary, Mother of Jesus. Since each unit has a different collection development policy and scope, each area also had different types of material that it sought to collect using Archive-It. In the case of University Archives and Special Collections, the unit had an urgent need to capture institutional content, including the university website and social media platforms. Increasingly, content that was previously published in paper format and easy for the University Archives to acquire is now created and distributed online. This unit's mission of capturing relevant institutional content was a primary factor in the University of Dayton Libraries' adoption of a web archiving tool, though it enhanced the collecting ability for all of the special collection departments.

In the case of the U.S. Catholic Collection, the Collections Librarian/Archivist created a collection of ten Catholic blogs, with their creators' permission. This collection supports the unit's mission of documenting Catholic life in the United States. It is crawled quarterly for archiving in Archive-It. The Marian Library chose to capture the organizational website of the Mariological Society of America, a theological organization whose

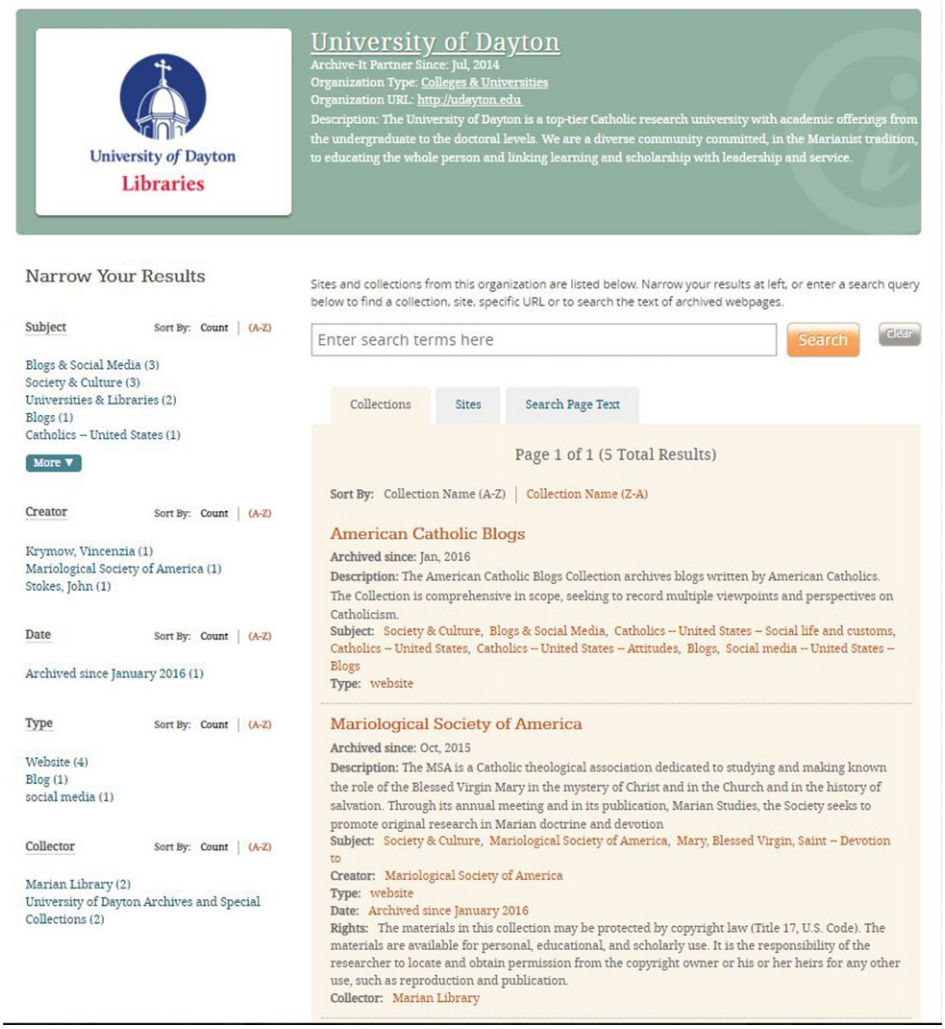

Figure 1. The University of Dayton's Archive-It collection page.

paper materials are maintained at the University of Dayton. The Marian Library Archivist also captured a blog related to a significant collection within the archives. All of these collections complement and supplement existing physical collections.

From the moment the Libraries' web archiving began, the public could access the archived web content through the Archive-It collection page. However, the material was not being actively promoted and the page was not linked from the Libraries website. The decision not to link to the collection page was originally made so that the archivists could gain more experience with collecting archived websites before promoting it. Description was also an issue, as there was no consensus within the Libraries or in the field of web archiving regarding metadata standards and appropriate terminology.

The archivists felt it was important to have some experience capturing the content before promoting it to users, but personnel changes pushed the project back even further. In late 2017, the Libraries hired a Discovery Services Librarian, a new position with job duties specifically relating to the discovery layer. Following that appointment, the University Archivist reached out to the Discovery Services Librarian to discuss potential ways to integrate the Archive-It content with other collections.

Literature review

Since 2011, the National Digital Stewardship Alliance (NDSA) – a consortium of universities, government agencies, professional associations, and other entities with an interest in digital preservation – has conducted four surveys of institutions that have web archiving programs to learn more about the current state of the field. The surveys, conducted in 2011, 2013, 2016, and 2017, have gathered essential data regarding the types of organizations with web archiving programs; the state of their programs; the tools and services they use; their access and discovery systems; and their program policies. The most recent survey from 2017 found that 60.5% of organizations engaging in web archiving were colleges or universities (National Digital Stewardship Alliance, 2018). Public libraries made up 12.6% of the total, a large increase from 3% in previous surveys. The increase in public libraries was likely due to the Internet Archive's "Community Webs" program, which offered grants to public libraries to begin web archiving programs. Other institutions with web archiving programs include federal and state governments (13.5% combined); and historical societies, commercial organizations, consortia, K-12 schools, museums, and "other" (13.4% combined).

Literature on this topic similarly shows that cultural institutions of all types have engaged in web archiving over the last two decades. Wallwork, Sutton, Whitver, and Fussell (2012) described the Tennessee State Library and Archives (TSLA)' creation of a web archive that preserves online state government documents and websites. The TSLA also created two unique access points for patrons to find their web archive: catalog records for the web archive and a "Tennessee Government Web Archive" webpage. Similarly, art museums and other art institutions have performed a great deal of web archiving in order to capture art resources that are now rarely printed, but almost exclusively found online. The National Museum of Women in the Arts started a web archive in order to preserve web ephemera related to women artists (Slania, 2013). Likewise, the New York Art Resources Consortium (NYARC) has archived various types of online art resources, including *catalogs raisonnés*, artists' websites, auction house

websites, and scholarship on the restitution of lost or looted art (Duncan, 2015; Duncan & Blumenthal, 2016). College and university archives have also started web archiving initiatives, frequently focusing on preserving their institutional websites, digital records, and social media (Antracoli, Duckworth, Silva, & Yarmey, 2014; Fansler, Gilbertson, & Petersen, 2014; Heil & Jin, 2017). Other universities have begun using web archiving as a collection development tool, creating web archives on topics such as the Ukraine conflict of 2014-2015 (Pendse, 2016) and Russian and Eastern European LGBT websites (Pendse, 2014).

Examining the digital preservation practices of Midwestern universities, Gorzalski (2018) found that 39% of these schools were preserving their institutional websites; 27% were preserving social media; and 16% were preserving personal websites. Ten institutions (out of 33) reported using Archive-It for acquisition, ingest, and preservation, and five institutions (out of 34) reported using Archive-It for access. Littman et al. (2018), in describing their creation and use of a "Social Feed Manager" (SFM) program for collecting data from Twitter's API, have advocated that social media archiving practices and tools should align with those of web archiving.

Another survey on the current state of web archiving found that the number of web archiving initiatives worldwide rose from 42 in 2010 to 68 in 2014, an increase of 61.9% (Costa, Gomes, & Silva, 2017). The number of countries hosting these initiatives increased from 26 in 2010 to 33 in 2014, an increase of 26.9%. Despite these advances, the 17 petabytes of data captured through web archiving initiatives are still just a small fraction of the total amount of data available online.

As web archiving has become more common, archivists and librarians have begun to study more closely various issues associated with this process. The OCLC report *Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group* (2018) found that metadata creation presents a significant challenge for web archiving practitioners. These findings echoed those of the NDSA, which reported in 2017 that metadata and description were the second-most-common area of "least progress in the last two years" (44% of respondents). The OCLC report stemmed from two surveys, one of the end users of web archiving content and one of the web archiving practitioners. These identified a lack of consistent metadata as the largest challenge facing the web archiving community as a whole. As a result, OCLC Research established the Web Archiving Metadata (WAM) Working Group to develop recommendations on metadata that will facilitate the discovery of archived websites. WAM has sought to offer guidelines on metadata creation but not strict mandates; it has intended its recommendations to fit the needs of any type or size of web archive. The OCLC report also includes a data dictionary

of metadata elements most important for description of web archives, and suggestions for future research on topics such as technical and preservation metadata, discovery layers, machine-actionable description, multiple levels of description, and MARC record types.

The 2017 NDSA survey also found that access and use of web archives remains problematic for many institutions. The most common area of "least progress" was access, use, and reuse of web archives collections (49% of respondents). Additionally, only 18% of respondents reported that they knew that their web archives were actively being used by researchers. Thirty-three percent said their web archives were not being used, and 49% reported that they did not know whether any use was occurring. These numbers showed little change from the 2016 survey.

Integration of web archives into discovery layers can increase the accessibility of these collections and encourage greater use of them. Several discovery layers use the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) to make metadata available from multiple resources in one place, increasing access for researchers (Corrado, 2018). There are many discovery layers available to libraries, including EBSCO Discovery Service (EDS), ProQuest's Summon and Primo, and OCLC's WorldCat Discovery Service, as well as open source options. Each of these products is capable of integrating Archive-It data. For example, NYARC uses an OpenSearch API to pull records from its Archive-It collections into Primo. It also creates catalog records for all of the websites that it archives, which display as part of the normal results set in Primo. Similarly, several other Archive-It subscribers have built custom portals to integrate their web content into their websites (Lohndorf, 2018).

Another issue in web archiving, which illustrates the technical challenges but is beyond the scope of this article, concerns the accuracy or completeness of webpage captures performed by web archiving software. Currently, these programs cannot always render with total accuracy the exact way a website looked and worked as it did in the past. With these limitations in mind, Brunelle, Kelly, SalahEldeen, Weigle, and Nelson (2015) proposed a new system that would measure the relative importance of web page elements. Their "damage rating algorithm" uses not just numerical calculations, but also human perceptions of the "completeness" of a webpage to measure how well it has been captured by web archiving tools.

Methods

The University of Dayton Libraries began to investigate the possibility of indexing Archive-It collections within UDiscover, the locally branded name of University of Dayton's EDS instance, in January 2018. The classic catalog, the consortial catalog (OhioLINK), and database holdings are indexed

within UDiscover. eCommons, the University's institutional repository (IR) hosted on Digital Commons, is set up as a custom collection and also indexed within UDiscover. Users can access UDiscover via the custom search box on the Libraries' homepage, research guides (LibGuides), and databases directory. A goal of the Libraries' 2017-2020 strategic plan was to "improve discovery and delivery of library resources" and the project to include Archive-It collections within the discovery service would help the library realize this aspect of the plan (University of Dayton Libraries, 2017). The project would positively impact user experience, while addressing the problem of silos of information.

A cross-departmental team within the Libraries was formed to work on the process of integrating the Archive-It collection into UDiscover. This team included four archivists, the Discovery Services Librarian, the Cataloging Coordinator, and the Digital Projects Manager. Initial communications occurred with EBSCO's Senior Software as a Service Catalog Specialist to inquire whether other libraries had integrated their Archive-It collections into EDS. Although some libraries had shown interest, none had completed the process. The team also placed inquiries on the EDS and Archive-It listservs, but no responses were received.

The team contacted NYARC about its indexing of Archive-It collections within Primo, using an OpenSearch API. In NYARC's setup, there was a 15-second lag while loading the web archive collections into the discovery search results, but its robust process, performed by a dedicated Web Archiving Coordinator and staff, provided a solid case for the team to proceed in creating a process to implement at our institution.

Using an API, as NYARC had done, was not a valid option due to the lack of systems and technological support at the University level. Upon further review of EDS documentation, the decision was made to create a custom IR in order to index the Archive-It collections. In addition, the eCommons collection had been set up previously as a custom IR, with no outlying issues. To begin the process, an "Institutional Repository Database Questionnaire" was submitted to EBSCO. The questionnaire asks for a harvesting schedule, the metadata format (Dublin Core), and specifications for the data and collections. In order for EBSCO to harvest the full dataset, an OAI-PMH URL must be provided. This URL was retrieved from the administrative end of the Archive-It account. The questionnaire was submitted to EBSCO's IR team in March 2018 so that it could work on the design, development, quality assurance testing, and production. Typically, EBSCO estimates 4–6 weeks for collection building, but since the collection contained only 30 records at the time, the process was completed more quickly. The new IR, Archive-It @ UD, was live and populating in UDiscover search results within 10 days.

Publication Type Lookup Table

The <TypeSource> mapped and Icon are fixed fields that are used to reflect your publication and document types present in your collections. This data is commonly located in <dc:type>, <dc:format> or <setName> tags within your data. Please review the list of viable publication types that EBSCO supports within its database and create a key in the table below to identify how your publication types should be represented. Add or delete rows as needed. If a document type in your data is not present in this lookup table, the value will be displayed as is in the database.

<dc:type> Type Source Field Content	<TypeSource> mapped	Icon (fixed per pub type)
Social Media	Electronic Resource	Electronic Resource
Facebook page	Electronic Resource	Electronic Resource
Instagram page	Electronic Resource	Electronic Resource
Twitter page	Electronic Resource	Electronic Resource
Blog	Website	Electronic Resource
Website	Website	Electronic Resource
Video	Video	Video

Figure 2. EBSCO Institutional Repository Publication Type Lookup Table.

Mariological Society of America

Collection:	Mariological Society of America
Access URL:	http://wayback.archive-it.org/6426/*/http://mariologicalsociety.com/
Accession Number:	aiud.75A46B41
Database:	Archive-It @

Figure 3. Archive-It @ UD record before metadata cleanup.

When reviewing the Archive-It @ UD collection, it was discovered that the source type icons and the accompanying text being displayed were inaccurate. All records defaulted to academic journal source types. This was due to not submitting an Institutional Repository Publication Type Lookup Table during the original collection building. In response to this discovery, the team agreed to a defined set of vocabulary for dc:type and the archivists updated the metadata records within Archive-It. Figure 2 shows the Institutional Repository Publication Type Lookup Table submitted to EBSCO that contained appropriate mapping for the type source and icons. Once submitted, EBSCO ran a full harvest of the Archive-It collection to implement the post-release change. This generated updated publication type icons in the records, as well as the correct publication type selections within the "Source Type" facet.

Additional editing of the metadata fields in Archive-It was necessary after examining the display of the collection in UDiscover. With three different collection owners, as well as a staffing change that occurred during the process, it was important to ensure each seed's metadata was standardized. Some records were missing titles and authors at the seed level, while others needed subject headings or descriptions. Figure 3 shows an example

Figure 4. Archive-It @ UD record after metadata cleanup.

of an original metadata record in UDiscover before the team set required fields.

Although EBSCO provides quality testing to ensure records are harvested and retrieved in search results, they do not review the metadata in the records. Based on the OCLC report *Descriptive Metadata for Web Archiving*…, the Coordinator of Cataloging created local guidelines that dictated which elements should be required at the seed and collection level in Archive-It in order to maintain consistency and standardization. The team reviewed each Archive-It @ UD record in UDiscover to note which metadata fields needed to be added or updated based upon the newly created guidelines and data dictionary. The local guidelines dictate that the "Collector", "Description", "Subject", "Title", "Type", and "URL" fields are required when creating a new seed in Archive-It. Figure 4 shows the same record as Figure 3, after the metadata was updated. The updated metadata fields for this record include the addition of "Collector", "Description", and "Subject". In addition, the frequency of the EBSCO harvesting schedule was updated from monthly to weekly so that the metadata changes would appear sooner.

Results

Collection and seed metadata fields were updated in Archive-It throughout the summer of 2018, with a plan to promote the resource in the fall. In addition to the Archive-It @ UD collection in UDiscover, additional access points were added, including a MARC record in the classic catalog, a link in the listing of databases on the Libraries website, and a database asset link in the LibGuides site. A web archiving LibGuide was created to explain the vision, scope, and guidelines of the program. When a user enters relevant keywords in UDiscover, a link to this guide is automatically displayed.

After the access points went live, the Libraries' Coordinator of Marketing and Engagement assisted with reaching out to the campus community. A post on the Archive-It collections was published to the Libraries' blog, which reached over 19,000 subscribers through the University of Dayton's e-newsletter. The blog was also promoted on the Libraries' Facebook page.

Information about Archive-It was shared with all University social media managers and added to the content management system training site for University website managers. They are encouraged to use the "Save Page Now" feature on the Internet Archive's Wayback Machine website before they edit and delete pages. This process adds to the archived snapshots within the collection without using any additional Archive-It subscription data.

A professional development event for University of Dayton librarians and library staff was held to share more information about web archiving and the Libraries' Archive-It collections. This was helpful to the Research and Instruction, and Access Services librarians and staff who interact with users daily. Attendees were also engaged in conversations concerning the importance of web archiving and reflected on how they can use these tools in both their professional and personal lives.

Usage data and analytics are currently tracked via Google and EBSCO. The administrative end of Archive-It allows a Google Analytics Tracking ID to be added to the collection. From this, a variety of analytics on audience, behavior, and traffic can be gathered. The Libraries were already tracking usage of the website and UDiscover with Google Analytics so it was a natural fit to add Archive-It. The Google Analytics "Acquisition Overview" chart in Figure 5 compares traffic sources from October 2018 and September 2018. Visitors from social media increased in October due to the Library's Facebook posting on Archive-It, which linked out to the Collections page.

A review of user behavior during the six-month period from October 2018 to March 2019, illustrated in Figure 6, shows traffic to the Archive-It Collection page has remained steady, with 50–60 users per month. User traffic to the Collection page is coming from a variety of channels, including the classic catalog, LibGuides, the Libraries website, eCommons, and Wikipedia.

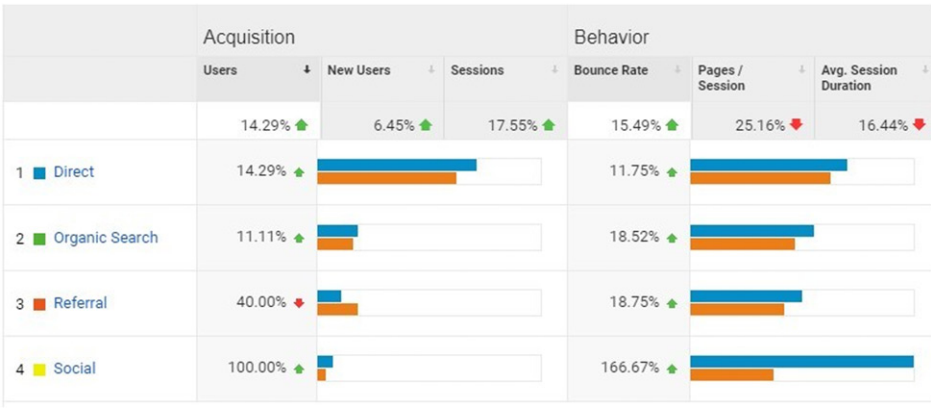

Figure 5. Google Analytics Acquisition Overview Chart comparing September 2018 (blue/dark) with October 2018 (orange/light).

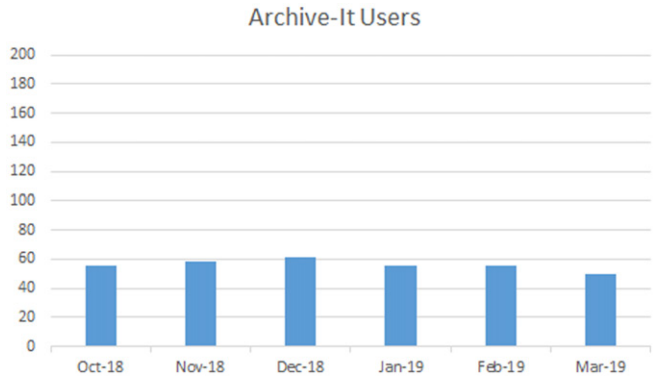

Figure 6. Number of users of the University of Dayton's Archive-It Collection page, October 2018–March 2019.

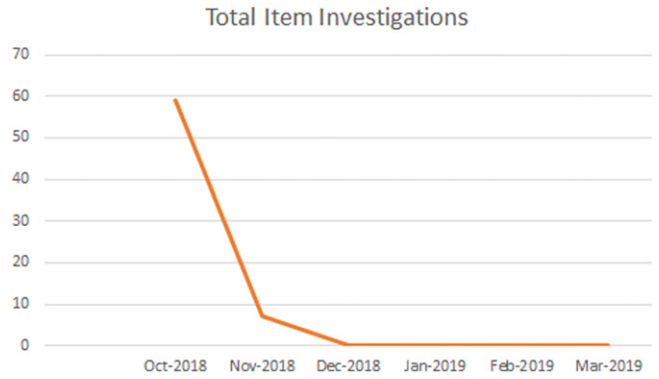

Figure 7. Archive-It @ UD item investigations, October 2018–March 2019.

There are a few options in EBSCOAdmin to run targeted usage reports of custom collections and repositories, including the standard database usage report, COUNTER 4 Database Report 1, COUNTER 5 Database Master Report, and COUNTER 5 Database Search and Item Usage. The latter counts the number of item investigations, meaning the number of times that users accessed a content item or information related to a content item. The COUNTER 5 Database Master Report in Figure 7 shows that October 2018 had the highest number of total item investigations. This usage can be attributed to the team's frequent review of the Archive-It @ UD collection records and promotional activities at that time. In comparison, item investigations in the following months dropped to zero. The use of collections may not be high, but it is important to make them accessible to ensure that they are available at the point of need. These items have significant potential reach, appearing in search results in UDiscover an average of 33,947 times per month. Usage statistics and analytics are collected and reviewed on an annual basis, which impacts future promotional efforts.

Conclusion

Success of this project is impossible to define by a quantitative analysis of usage statistics alone, since no data were collected prior to integration. Instead, success can be measured through the impact that the project has had on relationships internal and external to the library. There were few collaborative projects between archivists and technical services librarians, and each group had a general lack of understanding about the tools and products that the other group was using.

The professional development event offered by the team for Libraries faculty and staff was particularly worthwhile. While several attendees had heard of the Internet Archive's Wayback Machine, most did not know that the Libraries had a subscription to Archive-It or that the archivists were actively collecting websites. Librarians from the Research and Instruction teams were in attendance, and future discussions may include ways to use the web content for reference queries or in instruction sessions. Also, the Marian Library Archivist received suggestions of specific websites to consider crawling based on their unique Marian content after the professional development event.

Another outcome from the cross-departmental collaboration was the addition of the Discovery Services Librarian to the Libraries' existing Digital Preservation team. This team had recently issued a request for information (RFI) for digital preservation systems. Several preservation systems offer their own web crawling platform, similar to the service offered by Archive-It. The Discovery Services Librarian was able to ask the vendors detailed questions about their product's OAI-PMH capabilities to ensure that a switch to a new web archiving service would still allow the same workflow for accessing web archives from the discovery layer.

In addition to the internal relationships within the Libraries, the project has strengthened connections between several University departments and the University Archives and Special Collections. The training material about using the "Save Page Now" function of the Wayback Machine and the targeted promotion have increased the overall understanding of digital records. Some departments may not have previously recognized their website content as an electronic record worthy of saving in the University Archives. The archivists have seen this project open the door for further conversations about other types of digital records at the University of Dayton.

The six months of usage data present many avenues for future research. Figure 5 shows an increase in traffic from social media in October 2018 after the Libraries' blog post promoting the availability of web content. After that peak, usage remained steady, though with modest numbers. Evidently, audiences cannot be sustained with a one-time outreach event. The team did not set a specific goal in terms of usage since the intention was to increase

access. If increased use becomes an institutional priority, the data demonstrate that continued promotion of the content might increase user traffic.

Institutions considering ways to increase traffic to their web archives have several options for integrating access into other systems, as described by Lohndorf (2018). Based on the areas of success discussed above, integrating the Archive-It collections into UDiscover was the right solution at the University of Dayton. Reviewing the OCLC report *Descriptive Metadata for Web Archiving…* and developing local metadata guidelines are good first steps for any institutions seeking ways to increase access to their web archives. Institutions should also evaluate their available resources, including existing discovery and web archiving platforms, and internal IT support, before choosing their preferred method.

About the author

Christina Beis is the Discovery Services Librarian at the University of Dayton Roesch Library where she manages the life cycle of electronic resources, including access, discovery, assessment, and problem resolution. Her research interests include electronic resource workflows, usability, and emerging technologies.

Kayla Harris is the Archivist for the Marian Library at the University of Dayton, where she manages archival collections documenting devotion to Mary, Mother of Jesus. Kayla's areas of interest include teaching with special collections, and creating engaging digital projects.

Stephanie Shreffler is the Collections Librarian/Archivist for the U.S. Catholic Special Collection at the University of Dayton. Her research interests include classroom use of special collections, and digital preservation.

ORCID

Christina A. Beis http://orcid.org/0000-0003-4749-1292
Kayla Nicole Harris http://orcid.org/0000-0002-1672-3022
Stephanie L. Shreffler http://orcid.org/0000-0002-2147-7495

References

Antracoli, A., Duckworth, S., Silva, J., & Yarmey, K. (2014). Capture all the URLs: First steps in web archiving. *Pennsylvania Libraries: Research & Practice*, *2*(2), 155–170. doi: 10.5195/PALRAP.2014.67

Brunelle, J., Kelly, M., SalahEldeen, H., Weigle, M., & Nelson, M. (2015). Not all mementos are created equal: Measuring the impact of missing resources. *International Journal on Digital Libraries*, *16*(3–4), 283–301. doi:10.1007/s00799-015-0150-6

Corrado, E. M. (2018). Discovery products and the Open Archives Initiative Protocol for Metadata Harvesting. *International Information & Library Review*, *50*(1), 47–53. doi: 10.1080/10572317.2017.1422905

Costa, M., Gomes, D., & Silva, M. (2017). The evolution of web archiving. *International Journal on Digital Libraries*, *18*(3), 191–205. doi:10.1007/s00799-016-0171-9

Duncan, S. (2015). Preserving born-digital catalogues raisonnés: Web archiving at the New York Art Resources Consortium (NYARC). *Art Libraries Journal*, *40*(2), 50–55. doi: 10.1017/S0307472200000213

Duncan, S., & Blumenthal, K.-R. (2016). A collaborative model for web archiving ephemeral art resources at the New York Art Resources Consortium (NYARC). *Art Libraries Journal*, *41*(2), 116–126. doi:10.1017/alj.2016.12

Fansler, C., Gilbertson, K., & Petersen, R. (2014). The missing link: Observations on the evolution of a web archive. *Journal for the Society of North Carolina Archivists*, *11*(1), 46&59. Retrieved from http://www.ncarchivists.org/wp-content/uploads/2017/08/JSNCA_2014_Vol11No1_Fansler-Gilbertson-Petersen.pdf

Gorzalski, M. (2018). Digital preservation practices among midwestern four-year public colleges and universities. *Archival Issues: Journal of the Midwest Archives Conference*, *39*(1), 26–52. Retrieved from https://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1085&context=morris_articles

Heil, J. M., & Jin, S. (2017). Preserving seeds of knowledge: A web archiving case study. *Information Management Journal*, *51*(3), 20. Retrieved from http://imm.arma.org/display_article.php?id=2786569&view=408769

Littman, J., Chudnov, D., Kerchner, D., Peterson, C., Tan, Y., Trent, R., … Wrubel, L. (2018). API-based social media collecting as a form of web archiving. *International Journal on Digital Libraries*, *19*(1), 21–38. doi:10.1007/s00799-016-0201-7

Lohndorf, J. (2018, September 20). Archive-It access integrations. [blog post]. Retrieved from https://support.archive-it.org/hc/en-us/articles/360001231286-Archive-It-Access-Integrations

Matisse, N. (2018, October 7). The Internet's keepers? "Some call us hoarders - I like to say we're archivists". *Ars Technica*. Retrieved from https://arstechnica.com/gaming/2018/10/the-internets-keepers-some-call-us-hoarders-i-like-to-say-were-archivists/

National Digital Stewardship Alliance. (2018). *Web archiving in the United States: A 2017 survey*. https://doi.org/10.17605/osf.io/3qh6n

Pendse, L. R. (2014). Archiving the Russian and East European lesbian, gay, bisexual, and transgender web, 2013: A pilot project. *Slavic & East European Information Resources*, *15*(3), 182–196. doi:10.1080/15228886.2014.930973

Pendse, L. R. (2016). Analyzing the contemporary Czech literary web. *Slavic & East European Information Resources*, *17*(4), 285–297. doi:10.1080/15228886.2016.1246322

OCLC Research. (2018). *Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group*. Dublin, OH: OCLC Research. https://doi.org/10.25333/C3005C

Slania, H. (2013). Online art ephemera: Web archiving at the National Museum of Women in the Arts. *Art Documentation: Bulletin of the Art Libraries Society of North America*, *32*(1), 112–126. doi:10.1086/669993

University of Dayton Libraries. (2017, February 23). University of Dayton Libraries Strategic Plan 2017-2020 [blog post]. Retrieved from https://udayton.edu/blogs/libraries/2017-02-23-strategic-plan.php

Wallwork, T., Sutton, S., Whitver, A., & Fussell, B. (2012). Archive-It at the Tennessee State Library and Archives. *Tennessee Libraries*, *62*(2), 1–3. Retrieved from https://www.tnla.org/page/492