Digital preservation services at digital scholarship centers

Mikala Narlock [a,*], Daniel Johnson [b], Julie Vecchio, Assistant Director [c]

[a] *Digital Collection Strategy Librarian, Hesburgh Libraries, University of Notre Dame, United States of America*
[b] *English, Film, Television, and Theatre; Digital Humanities Librarian, Hesburgh Libraries, University of Notre Dame, United States of America*
[c] *Navari Family Center for Digital Scholarship, Hesburgh Libraries, University of Notre Dame, United States of America*

A R T I C L E   I N F O

A B S T R A C T

As academic library support services for digital scholarship activities continue to expand and evolve, large volumes of digital outputs have been created by, and in collaboration with, library and information professionals who are affiliated with digital scholarship centers. Drawing on a literature review and a 2018 pilot study of digital preservation services in digital scholarship centers, we propose future directions for investigation of preservation services for digital scholarship and projects.

Introduction

The proliferation of digital infrastructure, tools, and data sources has facilitated new types of academic exploration and created opportunities for novel collaborations with academic library specialty research support services, such as digital scholarship centers (DSCs) (e.g., Bryson et al., 2011; Johnson & Dehmlow, 2019). DSCs are described as a "service model in academic libraries that bring faculty and student scholars, technologists, and librarians together to collaboratively develop digital projects supporting scholarship and research" (Tzoc, 2016), and for the purposes of this research, digital scholarship is construed broadly as the use of digital evidence and methods, digital publishing, digital curation and preservation, and digital use and reuse of scholarship, regardless of discipline (Rumsey, 2011). Academic library support for digital scholarship encompasses a broad range of services, including teaching, consultation, outreach, the provision of access to technologies and data sources for creating and sharing new knowledge, and the creation and management of technology-enhanced spaces (e.g., Lippincott, 2017; Locke, 2017). As digital scholarship activities and outputs increase over time, the need for careful planning for the curation and long-term preservation of digital objects and projects is of critical importance (Owens, 2018). We explore the intersection of academic library digital scholarship centers with digital curation and preservation activities through the lens of a literature review and a 2018 pilot survey, seeking to address the following topics:

1. How do digital scholarship centers provide digital preservation information to their users?

2. What digital preservation support is provided by digital scholarship centers to their users?

3. What kinds of relationships and interactions can we observe between academic libraries, DSCs, and digital preservation activities?

Literature review

The expansive growth of digital scholarship work—along with a concomitant need for data—has resulted in strengthened connections between library and information professionals and digital scholars, especially digital humanists (Johnson & Dehmlow, 2019; Millson-Martula & Gunn, 2017; Sula, 2013). In particular, digital curation and preservation have been identified as ideal opportunities for collaboration between scholars, librarians, and information professionals, as library organizations tend to focus on lifecycle management with an emphasis on curation and preservation (Lippincott, 2017). While researchers may lack specific training for research data curation or experience with building and applying robust preservation policies, library and information professionals have been developing and utilizing these skills for decades (Poole & Garwood, 2018).

Tenopir, Birch, and Allard (2012, 5) argue that there are "powerful reasons for librarians to explore how their academic libraries can better satisfy the needs of researchers in the new data-intensive research atmosphere," including the curation of research data to facilitate discovery, and advocacy for effective preservation. As Walters and Skinner (2011) note, when "the library embeds the curation and preservation infrastructure and knowledge within its own staffing and digital framework and provides stable, trustworthy, and affordable services to

---

* Corresponding author.
E-mail addresses: mnarlock@nd.edu (M. Narlock), djohns27@nd.edu (D. Johnson), jvecchio@nd.edu (J. Vecchio).

its campus, the library as an institution becomes more secure and influential within its campus setting." (24) However, the results of Tenopir et al.'s, 2012 survey suggest that, while academic libraries and librarians are capable of providing research data services and support, there are often serious limitations in funding, particularly for staffing and repository maintenance. Since then, academic libraries have diverted more resources for research data services (particularly at R1 institutions), staff development, and additional support positions (Tenopir et al., 2019). This increased support coincides with increased collaborative efforts between libraries and DSCs, which academic libraries have leveraged as an opportunity to advocate for their position, funding, and new roles (Cox, 2016). This work has also resulted in increased development of tools to support data and digital project curation, including efforts such as the Preservation Quality Tool (PresQT) and Emulation as a Service Infrastructure (EaaSI), which help harvest and curate data and metadata, and ensure that, regardless of format, data will be accessible into the future, easing the burden on both information professionals and repository managers.

The importance of digital curation, and specifically lifecycle management, has been written about extensively within the context of specific types of disciplines, as well as writ large. In the humanities, digital curation has been supported by grant-funded projects such as the University of Pittsburgh's "Sustaining DH" NEH Institute for Advanced Topics in the Digital Humanities (https://sites.haa.pitt.edu/sustainabilityinstitute/). The Institute educated librarians and departmental faculty alike on a new "Socio-Technical Sustainability Roadmap," a framework to assist in "the seemingly daunting task of sustaining … web-based, user-facing, digital humanities project over time" (https://sites.haa.pitt.edu/sustainabilityroadmap/getting-started/). Similar efforts include "The Endings Project," funded by the Social Sciences and Humanities Research Council of Canada (https://projectendings.github.io/), Katherina Fostano and Laura K. Morreale's "Digital Documentation Process" for DH scholarship (https://digitalhumanitiesddp.com/), and the Mellon-supported "Digits Project," which promises to "conduct an environmental scan of the use of software containers in research and publication, as well as a fact-finding mission on the infrastructural needs of scholars who are currently producing non-standard digital research" (https://digits.pub/about/).

Social science data are among the oldest digital media: beginning in the late 1800s, US census data were converted to a digital format for analysis by—what was at the time—brand-new tabulating machines (Gutmann et al., 2009). Text-mining and artificial intelligence technologies available today are further extending the variety of data available for exploration through social science methodologies, shifting "the evidence base of social science" (Walters & Skinner, 2011). Social science data pose complex and unique challenges for data curation and preservation: documentation may be lacking or inaccessible, data ownership may be in question, data may have rigorous privacy/confidentiality requirements, and data format persistence may be problematic (ICPSR, 2012; Lyle et al., 2014). Repositories—both institutional and disciplinary—are vital to the preservation of social science research assets and outputs, but are bound by their own unique missions and policies. Collaborative projects such as the Data Preservation Alliance for the Social Sciences (Data-PASS: http://www.data-pass.org/) leverage the resources of multiple institutions in support of the identification, acquisition, and curation of social science data that have been deemed "at risk," whether from legacy research sources or from ongoing or future work (Gutmann et al., 2009). Academic library and information professionals—whether affiliated with DSCs or not—play a variety of critical roles in the preservation of social science data, ranging from acquisition, to educational and outreach services, to hands-on curation work, to name a few (Tammaro et al., 2019; Xia & Wang, 2014).

The 'hard sciences' tend to produce data at a larger scale than the social sciences and humanities, especially that which is derived from niche software, tools, and highly-advanced equipment. Researchers and information professionals have actively been working to provide

persistent and long-term access to research data and other scholarly outputs. Since the early 2000's, librarians and information professionals have been advocating for and documenting research data curation (e.g., Gray et al., 2002), articulating the lifecycle of research data (e.g., Higgins, 2008), and carving space for information professionals to assist in the curation process. Data curators, discipline experts, and even private companies have developed numerous tools to help scholars and repository managers preserve content and provide consistent access to data and digital objects. The proliferation of disciplinary, institutional, and general repositories for researchers, as well as curatorial tools like wholeTALE, facilitate not only data reuse and reproducibility, but also curation and long-term accessibility to the data. In recent years, the rise of FAIR data (Findable, Accessible, Interoperable, and Reusable; Wilkinson et al., 2016), increasing funder mandates and required data management plans (DMP), and hands-on data sharing workshops and hackathons (e.g., Hildreth & Meyers, 2020) have resulted in an increased awareness around the intricacies of preserving research data and the need to define domain-specific requirements.

Despite the prevalence of digital scholarship activities across academic disciplines, preservation remains a persistent challenge bedeviled by uncertain expectations, uneven work distribution, and inadequate sustainability planning, among other issues. Atkins (2013) found that most organizations, when lacking a dedicated digital preservation program, often left the task of preservation to the library. Li et al. (2020) observed a similar desire for help with managing research data at Wuhan University Library, but found in a quantitative survey that researchers "do not entirely believe librarians can be of significant help in managing research projects, providing data curation and sharing support," leading them to suggest that libraries should "promote and advertise their effort and abilities" (9.) Libraries, however, may lack the funding or technical infrastructure needed to support digital projects adequately in the long term (Owens, 2018). Moreover, given that effective digital preservation and consistent, long-term access to the content requires intense curatorial support, librarians, specifically subject selectors and disciplinary curators, are in the best position to provide feedback on digital scholarship projects (Tallman & Work, 2018). Robert Montoya (2017, 221) even argues that a new category of "*boundary staff* specifically charged with maintaining … boundary infrastructures and negotiating mismatched practices between departments" is needed to break out of silos and integrate library strengths with cross-disciplinary projects.

Regardless of where a digital object or project originates or concludes, the stakes for digital preservation are high, and project partners benefit from sharing the responsibility and privilege of applying digital preservation considerations to their work. Indeed, increasing the pool of stakeholders should increase preservation options, helping to alleviate the burden of hidden labor on a small group of individuals while also avoiding the temptation to overfit all projects to a one-size-fits-all preservation solution. DSCs, in turn, stand to benefit by learning how their peers are engaging stakeholders in this important endeavor.

Pilot survey

For additional perspective on this landscape, we distributed a pilot survey via list-serv in order to investigate how digital scholarship centers within higher education institutions in the United States currently engage with their stakeholders on digital preservation. In total, the survey received forty-seven (47) responses. Respondents who left all answers blank were eliminated. Duplicate responses were received from three institutions. If there was overlap between responses, the authors looked to see if responses were identical; if so, one entry was kept for the institution, and if not, both entries were removed. Two entries were removed as non-US institutions. In total, twenty-five (25) survey responses were used for analysis. For more information, please visit https://doi.org/10.17605/OSF.IO/3YJ8A.

A key limitation of this survey is the small number of responses

received relative to the number of invitations distributed through list-servs; the survey nevertheless provides an instructive starting place for continued exploration of digital preservation patron engagement activities at US digital scholarship centers. In following the format of the survey, the themes that emerged from our data have been divided into two categories: characteristics of the responding DSCs, and patterns of digital preservation practices.

Responding digital scholarship center overview

All responding centers indicated that they provide consultations to patrons ($n = 25$). Most responding DSCs indicated that they provide instruction ($n = 22$), cultivate a web presence ($n = 21$), and provide access to hardware and software for patron use ($n = 20$). Responding DSCs tended to have a broad range of expertise: while the particulars varied between DSCs, many indicated that they offer expertise in digital publishing, project management, data analysis, and metadata ($n = 22$, 20, 19, 19). Digital preservation was an area of expertise for over half of responding DSCs ($n = 16$), followed closely by institutional repository support ($n = 15$). Areas of DSC expertise may warrant additional exploration, specifically the emphasis on project management and data analysis and how they relate to preserving digital scholarship. The responses here could be indicative of a number of things, including but not limited to: a primary focus on active project development by responding DSCs, which are often on the cutting edge of research and research methods; the possibility that responding DSCs were collaborating with patrons on sustainable projects that need less preservation support; a prevalence of projects that had not yet reached a stage where preservation concerns are imminent; or perhaps a lack of interest in preservation among responding DSC patrons. Additional investigation into these motivations for prioritizing project management and data analysis could help guide future developments in DSC support for curating and preserving digital scholarship outputs.

Physically and organizationally, responding DSCs were linked to libraries, echoing the prevalent themes in the literature about the relationship between the two (Lippincott & Goldenberg-Hart, 2014). Most respondents noted that their DSC is located organizationally with the institution's library ($n = 19/25$, 76%), and, when asked about their roles and responsibilities within the DSC, approximately one third of respondents indicated that their primary role was that of "Librarian" ($n = 9/25$). A responding DSC's connection with an academic library was not associated with provision of digital preservation support by the responding DSC. This is an area that may warrant additional exploration: Given libraries' and archives' legacy of preservation and providing long-term access to materials, the library is the heir-apparent to preserving content created by and with the DSC, whether through curation, storage, metadata/descriptive practices, or other preservation activities. However, limited funding, overwhelmed staff, and DSCs' charge to stay at the forefront of digital scholarship may prohibit this collaboration.

Digital preservation practices of responding digital scholarship centers

In terms of audience for digital preservation support, the majority of responding DSCs ($n = 19$) indicated that they provide support for digital preservation to patrons, with the primary demographic overwhelmingly faculty-oriented and humanities-centric. This could be due to the wide definition of "digital scholarship center" employed by the survey, which included digital humanities centers under the digital scholarship center umbrella. Additional exploration of the core demographics of communities who engage with DSC services could be helpful for guiding the development of additional best practices for engaging users in digital preservation conversations.

Overwhelmingly, the digital preservation support provided by responding DSCs tended to take the form of consultations ($n = 19$), followed by instruction and outreach ($n = 8$). This suggests an opportunity for developing additional resources for the integration of reusable assets and frameworks into consultative and instructional sessions.

Future explorations and conclusion

The literature review points to ample opportunities for libraries to engage across disciplines in digital preservation, and warns of peril if they don't. Our pilot survey responses, though limited, suggest specific avenues of research, including the expansion of primary audiences for digital preservation outreach, the development of new (or implementation of existing) resources for engaging faculty and students in digital preservation activities compatible with the time limitations in outreach and consultation, and consideration of the implications of organizational placement of DSCs for the provision of digital preservation support to patrons.

As DSCs continue to evolve, academic library organizations should consider prioritizing digital preservation competencies in continuing education opportunities for their employees. According to King (2018), there are a number of skills useful for DSC faculty and staff, including technical abilities, but also more traditional librarian expertise, including preservation, institutional repository support, and metadata enhancement; however, "Librarians felt overwhelmingly that they needed more, better trained staff to meet this need and that they themselves were in need of skills, knowledge and credentials." (44) By providing these educational opportunities, funding, or other support to employees in addition to DSC patrons, libraries can continue to serve as active and collaborative partners in supporting the creation and preservation of digital objects and digital scholarship projects.

As a follow-up to this work, more detailed investigation into preservation, through activities such as semi-structured interviews with survey respondents, could provide even more specific information on how DSCs engage patrons. While the pilot survey provides a snapshot in time, the response categories were too broad to learn detailed information at the outset. Additional research could investigate how active subject selectors, curators, or other disciplinary liaisons are in supporting the curation and preservation of DSC projects. Relatedly, we would like to learn whether DSCs are providing rubrics or other tools to support curators in deciding what to preserve, and to see how many DSCs are embracing a benign neglect towards their projects, allowing them to gracefully decline.

Since the initial distribution of the survey, the landscape of higher education has changed drastically in the wake of the COVID-19 pandemic. Additional research could examine how remote work has impacted consultations and remote digital preservation work. Similarly, during the myriad social protests that occurred during the Summer of 2020, did DSCs engage in or support community archiving or preservation?

The results of this pilot survey and related research have uncovered more questions than answers. As libraries and DSCs contend with an ever-increasing proliferation of data and digital objects—especially when considering legacy digital projects from early-adopters in the 2000s—and budgets that remain constant at best, effective digital preservation relies on an active collaboration between partners. Knowing how best to support DSCs and library and information professionals in this endeavor ensures time and resources are spent effectively in providing long-term access to digital projects for future scholars. This work can and must be a collaborative effort between institutional and organizational units, and requires more investigation to understand just where to start.

References

Atkins, W. (2013). *Staffing for effective digital preservation: An NDSA report: Results of a survey of organizations preserving digital content.* National Digital Stewardship Alliance.

Bryson, T., Posner, M., St. Pierre, A., & Varner, S. (2011). Digital humanities, SPEC kit 326 (November 2011). https://publications.arl.org/Digital-Humanities-SPEC-Kit-326/.

Cox, J. (2016). Communicating new library roles to enable digital scholarship: A review article. *New Review of Academic Librarianship, 22*(2–3), 132–147. https://doi.org/10.1080/13614533.2016.1181665

EaaSI GitLab | Software Preservation Network (SPN). (n.d.). Retrieved from https://www.softwarepreservationnetwork.org/eaasi-gitlab/.

Gray, J., Szalay, A. S., Thakar, A. R., Stoughton, C., & van den Berg, J. (2002). In A. S. Szalay (Ed.), *Online scientific data curation, publication, and archiving* (pp. 103–107). https://doi.org/10.1117/12.461524

Gutmann, M. P., Abrahamson, M., Adams, M. O., Altman, M., Arms, C., Bollen, K., , … King, G., et al. (2009). From preserving the past to preserving the future: The data-PASS project and the challenges of preserving digital social science data. *Library Trends, 57*, 315–337. Internet.

Higgins, S. (2008). The DCC curation lifecycle model. *International Journal of Digital Curation, 3*(1), 134–140. https://doi.org/10.2218/ijdc.v3i1.48

Hildreth, M., & Meyers, N. (2020). *Final report: FAIR Hackathon workshop for mathematical and physical sciences research communities*. https://doi.org/10.7274/R0-RWPP-AS13

Inter-university Consortium for Political and Social Research (ICPSR). (2012). *Guide to Archiving Social Science Data for Institutional Repositories* (1st ed.) Ann Arbor, MI.

Johnson, D., & Dehmlow, M. (2019). Digital exhibits to digital humanities: Expanding the digital libraries portfolio. In *New top technologoies every librarian needs to know: A LITA guide* (p. 123).

King, M. (2018). Digital scholarship librarian: What skills and competences are needed to be a collaborative librarian. *International Information & Library Review, 50*, 40–46. https://doi.org/10.1080/10572317.2017.1422898

Li, B., Song, Y., Lu, X., & Zhou, L. (2020). Making the digital turn: Identifying the user requirements of digital scholarship services in university libraries. *The Journal of Academic Librarianship, 46*(2), Article 102135. https://doi.org/10.1016/j.acalib.2020.102135

Lippincott, J. K. (2017). Opening keynote: Fulfilling our mission in the digital age. *Digital Initiatives Symposium*, 17. Retrieved from https://digital.sandiego.edu/cgi/viewcontent.cgi?article=1131&context=symposium.

Lippincott, J. K., & Goldenberg-Hart, D. (2014). CNI workshop report. Digital scholarship centers: Trends and good practice. Retrieved from https://www.cni.org/wp-content/uploads/2014/11/CNI-Digitial-Schol.-Centers-report-2014.web_.pdf.

Locke, B. T. (2017). Digital humanities pedagogy as essential liberal education: A framework for curriculum development. *Digital Humanities Quarterly, 011*(3).

Lyle, J., Alter, G., & Green, A. (2014). *Partnering to curate and archive social science data. Research data management: Practical strategies for information professionals* (pp. 203–222).

Millson-Martula, C., & Gunn, K. (2017). The digital humanities: Implications for librarians, libraries, and librarianship. *College & Undergraduate Libraries, 24*(2–4), 135–139. https://doi.org/10.1080/10691316.2017.1387011

Montoya, R. D. (2017). Boundary objects/boundary staff: Supporting digital scholarship in academic libraries. *The Journal of Academic Librarianship, 43*(3), 216–223. https://doi.org/10.1016/j.acalib.2017.03.001

Owens, Trevor (2018). *The theory and craft of digital preservation*. John Hopkins University Press.

Poole, A. H., & Garwood, D. A. (2018). "Natural allies": Librarians, archivists, and big data in international digital humanities project work. *Journal of Documentation, 74*(4), 804–826. https://doi.org/10.1108/JD-10-2017-0137

Rumsey, A. S. (2011). *Scholarly communication institute 9: New-model scholarly communication: Road map for change*. Charlottesville, VA: University of Virginia Library.

Sula, C. A. (2013). Digital humanities and libraries: A conceptual model. *Journal of Library Administration, 53*(1), 10–26. https://doi.org/10.1080/01930826.2013.756680

Sustaining DH – An NEH Institute for Advanced Topics in the Digital Humanities. (n.d.). Retrieved from https://sites.haa.pitt.edu/sustainabilityinstitute/.

Tallman, N., & Work, L. (2018). Approaching Appraisal. In *International Conference on Digital Preservation (Vol. 2018)*.

Tammaro, A. M., Matusiak, K. K., Sposito, F. A., & Casarosa, V. (2019). Data curator's roles and responsibilities: An international perspective. *Libri, 69*(2), 89–104. https://doi.org/10.1515/libri-2018-0090

Team, T. E. P. (n.d.). The Endings Project. Retrieved from https://endings.uvic.ca/.

Tenopir, C., Allard, S., Baird, L., Sandusky, R., Lundeen, A., Hughes, D., & Pollock, D. (2019). Academic librarians and research data services: Attitudes and practices. *IT Lib: Information Technology and Libraries Journal*, (1). https://trace.tennessee.edu/utk_infosciepubs/99.

Tenopir, C., Birch, B., & Allard, S. (2012). Academic libraries and research data services: Current practices and plans for the future. In *An ACRL white paper*. https://trace.tennessee.edu/utk_dataone/20.

The Digital Documentation Process—The Digital Documentation Process. (n.d.). Retrieved from https://digitalhumanitiesddp.com/.

Tzoc, E. (2016). Libraries and faculty collaboration: Four digital scholarship examples. *Journal of Web Librarianship, 10*(2), 124–136. https://doi.org/10.1080/19322909.2016.1150229

Walters, T., & Skinner, K. (2011). *New roles for new times: Digital curation for preservation*. Association of Research Libraries. https://vtechworks.lib.vt.edu/handle/10919/10183.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18

Xia, J., & Wang, M. (2014). *Competencies and responsibilities of social science data librarians: An analysis of job descriptions | Xia | College & Research Libraries*. https://doi.org/10.5860/crl13-435