

Sustaining Digital Humanities Collections: Challenges and Community-Centred Strategies

Katrina Fenlon
College of Information Studies,
University of Maryland, College Park

Abstract

Since the advent of digital scholarship in the humanities, decades of extensive, distributed scholarly efforts have produced a digital scholarly record that is increasingly scattered, heterogeneous, and independent of curatorial institutions. Digital scholarship produces collections with unique scholarly and cultural value—collections that serve as hubs for collaboration and communication, engage broad audiences, and support new research. Yet, lacking systematic support for digital scholarship in libraries, digital humanities collections are facing a widespread crisis of sustainability. This paper provides outcomes of a multimodal study of sustainability challenges confronting digital collections in the humanities, characterizing institutional and community-oriented strategies for sustaining collections. Strategies that prioritize community engagement with collections and the maintenance of sociotechnical workflows suggest possibilities for novel approaches to collaborative, community-centred sustainability for digital humanities collections.

Submitted 17 December 2019 ~ *Accepted* 19 February 2020

Correspondence should be addressed to Katrina Fenlon, 4130 Campus Drive, Hornbake Library South rm 4105, College Park, MD 20742. Email: kfenlon@umd.edu

This paper was presented at International Digital Curation Conference IDCC20, Dublin, 17-19 February 2020

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: <http://www.ijdc.net/>

Copyright rests with the authors. This work is released under a Creative Commons Attribution Licence, version 4.0. For details please see <https://creativecommons.org/licenses/by/4.0/>



Introduction

Since the advent of digital scholarship in the humanities, decades of extensive, distributed scholarly efforts in collecting and digitization, datafication, modelling, encoding, scholarly editing, annotation, and the development of maps, games, simulations, and more, have resulted in a digital scholarly record that is increasingly scattered, heterogeneous, and independent of libraries and cultural institutions. The digital outputs of humanities research are increasingly media-rich, data-centric, interactive, and interlinked with external resources. They are also increasingly common; more than half of faculty report creating digital tools and collections, most intended for public use or to serve a disciplinary community of researchers (Maron and Pickle, 2014). Digital scholarship produces collections with unique scholarly and cultural value, both in their capacity to manifest scholarly interpretation and serve new research and reuse, and in their propensity to gather and represent digital primary source evidence that does not exist as such in mainstream memory institutions.

Yet the bulk of digital humanities collections are unsustainable. Outside of well-resourced digital humanities centres and libraries, there continues to be a systematic lack of support for digital scholarship after the phase of its initial creation. Even on campuses with established digital humanities centres, there are rarely end-to-end solutions in place for supporting digital scholarship from its conception to preservation, so that maintaining projects- which are built by scholars or research communities, often on bespoke infrastructures using short-term funding – has become a major problem for institutions (Maron and Pickle, 2014; Smithies et al., 2014). Library support for digital scholarship at every phase of its lifecycle is growing but remains profoundly inadequate overall to match the ongoing growth in digital scholarship or confront the existing accumulation of legacy collections.

This paper reports on a multimodal study of the sustainability challenges confronting digital collections in the humanities. Based on a set of interviews with practitioners in digital humanities centres and libraries, supplemented by an analysis of digital collections, this paper identifies the central challenges confronting the management of collections over time. This paper then characterizes strategies for sustaining collections, dwelling on one strategy in urgent need of increased research and understanding: that of community engagement with and reuse of digital collections in the humanities, with the goal of moving toward community-centred sustainment.

Background

One common mode of digital humanities production is the digital collection – often called *thematic research collection* (Palmer, 2004) or *digital archive* – which takes the form of a curated aggregation of primary sources along with materials and features designed to support research on a theme. “Collection” is used as a shorthand in this paper for a variety of digital projects and their outcomes, ranging from scholarly editions to linked data hubs, which gather primary sources or evidence derived from sources, and integrate those sources with annotation, contextual information, secondary sources, or functional and interactive elements in order to construct platforms for learning and research. Digital humanities collections serve as hubs for collaboration and communication, engage broad audiences, and generate new research (Palmer, 2004; Fenlon, 2017). While collections have long constituted a prominent mode of digital scholarship (Palmer, 2004; Flanders, 2014; Fenlon, 2017; Cooper and Rieger, 2018), they rarely gain integration into systems of digital curation or preservation in libraries and other curation institutions. Despite the fact that most fall well within scope of the preservation missions of libraries responsible for stewarding institutional research, digital humanities collections are facing a widespread crisis of sustainability.

Sustainability and preservation are uniquely problematic for digital humanities collections, for many reasons. Collections are often developed and maintained outside of the purview of dedicated memory institutions. They tend to be centred in scholarly communities, in the sense that scholars create and maintain collections for their own uses or the uses of their communities, with fluctuating resources, and usually without professional curatorial support. Because these collections tend to be funded on short cycles oriented toward technical innovation or experimentation and rapid development, they often rely on bespoke or fragile infrastructures. These collections are highly creator-dependent; they rarely endure beyond the interest and involvement of their initial creators, even when there are active communities of use. There is evidence of systemic confusion around the value of digital scholarship to academic institutions, and how institutions should understand ownership of highly collaborative and distributed projects (Maron and Pickle, 2014). And because collections function simultaneously as scholarly publications and as platforms for ongoing research, they confront a conceptual morass around what sustainability and preservation really mean for different kinds of digital scholarship in different contexts. More pragmatically, most academic libraries simply lack capacity to take in and sustain any more than a narrow swath of digital scholarship.

Sustainability is a term that has garnered widely varying definitions across the literatures of practice and research in cultural heritage, digital humanities, and digital curation. Most discussions of sustainability revolve around organizational resilience, long-term economic viability, and questions of institutional management (Eschenfelder et al., 2016). There is increasing recognition of the sociotechnical aspects of sustainability – of the need to maintain the collaborative processes and labour that serve to construct digital scholarship in combination with technical artifacts and processes (Langmead et al., 2018; Madsen and Hurst, 2018). This paper builds on sociotechnical approaches to sustainability, considering *sustainability* to mean the ability of a collection to remain viable over time, to responsively support the communities that create and use it, in whatever forms are useful, for as long as useful. In contrast to a paradigm of digital preservation focused on fixity, this definition of *sustainability* admits the need for collections to continue to change and grow. This definition also presumes that sustainability and preservation approaches exist on a spectrum, with no clear delineation between them.

Institutional efforts to sustain and preserve digital scholarship are commonly characterized by one or more of the following three main features: (1) maintenance and preservation efforts are solely or primarily assumed by digital humanities centres, where they exist; (2) where centres or preservation institutions (mostly libraries) offer long-term support for digital scholarship, that support is generally framed in terms of *service levels*; and (3) repository, publishing, and data management infrastructures are developed to increase the capacity of institutions to hold and maintain increasingly complex digital scholarship.

Digital humanities centres commonly serve as inadvertent, sometimes reluctant memory institutions. Depending on their capacity and their relationships with other entities, they make sporadic, often reactive investments into maintaining digital projects that they host. Some centres and labs have developed comprehensive strategies and policies to confront burgeoning maintenance needs (Smithies et al., 2019; Madsen and Hurst, 2018). Centres may possess a range of relationships with institutional libraries, ranging from complete independence to physical colocation and organizational ties. These relationships substantially affect the capacity of a center or lab to sustain digital scholarship over time (Prescott, 2015).

For both digital humanities centres and for libraries playing an active role in sustaining or preserving digital scholarship, the most common reported strategy involves the articulation and negotiation of a service model comprised of varying *service levels* or layers. Service levels are usually defined around the varying commitments a library or center agrees to make to maintain discrete kinds of components, significant properties, or levels of access to collections in response to identified functional requirements (e.g., Oltmanns et al., 2019; Madsen and Hurst, 2018; Goddard and Walde, 2017; Vinopal and McCormick, 2013; Sustaining Digital Scholarship, 2004). Service levels may be negotiated on a per-project basis to create formal agreements between digital humanities creators and libraries or centres, or they may constitute blanket institutional policies. For libraries, this layered service model in almost every case entails a

“handoff” of a collection – migration of the collection along with transfer of ownership or responsibility – from a research community to the library. At what point in the lifecycle of a project that handoff happens varies widely.

In addition to developing policies, a final common institutional strategy is the development or adoption of advanced technical infrastructure for the management, preservation, and publication of increasingly complex digital objects and collections. Emergent preservation repositories, publishing platforms, and collaborative research environments aim to capture and represent complex digital research objects, linked data, and primary source collections alongside and interleaved with traditional forms of scholarly publication (e.g., Sweeney et al., 2017; Almas, 2017; White et al., 2019; Fenlon, 2019). Digital humanities scholarship has generally resisted large-scale infrastructure for many reasons, including the high variation in user requirements across projects (Dombrowski, 2014), the non-scalability of digital humanities and digital curation (Rawson and Muñoz, 2019), and epistemological tensions with established and emergent cyberinfrastructure from other domains (Fenlon, 2019; Smithies et al., 2019).

Beyond institutionally centred strategies for digital humanities sustainability and preservation, there is a promising movement within cultural institutions toward shared stewardship and related models for partnering with communities to share the work of collection maintenance over time (e.g., Smithsonian, 2019). These models emerge from a substantial body of research in the archival community on post-custodial and participatory archives (Gilliland and Flinn, 2013; Caswell, 2014; Clement et al., 2013). While these efforts have largely focused on community archives rather than digital scholarship, they may offer a promising direction for digital collections more broadly.

Methods

This paper reports selected outcomes of a multimodal, qualitative study of thematic research collections as an emergent mode of digital scholarship in the humanities, along with challenges for libraries in supporting collections throughout their lifecycles. The study was conducted in three phases: (1) typological analysis of a large sample of collections (n=145), which characterized the range and defining features of collections; (2) qualitative content analysis of three exemplary collections to more deeply characterize the genre; and (3) a set of semi-structured interviews with nine practitioners, representatives of digital humanities centres and libraries, each with significant expertise in the creation and management of digital humanities collections. The goal of the interview phase of the study was to identify current practices in supporting thematic research collections, along with challenges and strategies for integrating collections into infrastructures of maintenance and preservation. This paper focuses on the outcomes of the interviews, which had the most bearing on questions of sustainability and preservation. However, a relevant outcome of the typology and content analysis phases of this study – which pertains to different modes of *contribution* of digital collections – is summarized in the first part of the “Challenges” section, below. For details on methods and findings of typology and content analysis, see Fenlon (2017).

The interview phase of this study addressed questions including: What are the challenges, for libraries and related scholarly-publishing entities, in supporting thematic research collections as a scholarly genre? How do library publishing programs and related scholarly-publishing entities support the creation and publication of thematic research collections, and what problems exist in meeting the needs of collection creators? How do libraries collect, represent, describe, preserve, and otherwise treat thematic research collections after publication, and what problems exist in meeting user needs? Sampling for the interview phase of the study was purposive. While the sample was small, participants were selected for their expertise in the creation and maintenance of thematic research collections, prioritizing the potential richness of expert response over any gains in generalizability that might be attained from a larger or random sample. Participants were selected to represent well-established centres and labs with a

long history of creating and maintaining digital collections, including the Center for Digital Research in the Humanities at the University of Nebraska-Lincoln, the Maryland Institute for Technology in the Humanities, the Roy Rosenzweig Center for History and New Media at George Mason University, and the Scholars' Lab and the Institute for Advanced Technology in the Humanities at the University of Virginia. All participants waived confidentiality for this study; nonetheless, the description of results below employs participant codes (in the form of "Participant X"), rather than names, to distinguish quotations by different participants. Where possible, interviews were conducted with more than one person from each institution. Two additional interviewees were selected for their extensive experience working with collections in addition to expertise in library administration. Interviews were coded using qualitative content analysis. The coding frame was built inductively, deriving themes from the transcripts in answer to the research questions.

The study admits several limitations beyond those that confront interview studies generally. This study focuses on the perspectives of collection creators within digital humanities centres (albeit, collection creators with significant expertise). Future work will need to integrate the perspectives of independent scholars, along with those of more and varied stakeholders in preservation institutions. Few libraries appear to systematically deal with thematic research collections post-publication, which makes empirical investigation of the possibilities difficult. For this reason, this study aims to be foundational rather than comprehensive or conclusive about the challenges confronting institutions.

Challenges to Sustaining Digital Humanities Collections

This study surfaced four main challenges confronting the sustainability and preservation of digital humanities collections: (1) Discontinuity between the essential interactivity of digital collections and the paradigm of artifactual preservation; (2) The importance and vulnerability of "connective tissue" within and between collections; (3) Ambiguity of institutional contexts and roles; and (4) Lack of infrastructure for collaborative humanities workflows. These challenges are grounded in and contextualized by an important observation about digital scholarship which emerged from the typological and content analysis phases of research: that the varying contributions of digital scholarship seriously complicate discourse around and practical approaches to sustainability and preservation.

Different collections aim to contribute to scholarship in different ways. This study identified different kinds of contributions that collections make to scholarship. While the contributions described here are by no means exhaustive, they exemplify epistemological differences that have a bearing on sustainability and preservation decisions. Based on the typological analysis and content analysis reported in Fenlon (2017), collections may be usefully differentiated by constellations of interrelated properties, such as a collection's purpose(s), a collection's theme or subject, the kinds and diversity of items in a collection, and how interrelationships among items in a collection are created through technical, narrative, and design elements. In fact, this study found that the combination of these properties may be boiled down to a deceptively simple question, with which to differentiate collections: What would it mean for a given collection to be *complete*? In other words, what idea of completeness – in the senses of wholeness, totality, or comprehensiveness – guides the development of the collection? The study identified three preliminary kinds of collections, each bent toward a different ideal of completeness:

- **Definitive source collections** aim to bring together an exhaustive set of definitive primary sources, to serve as an authoritative resource for scholarship. Sustainability and preservation efforts for such collections would likely centre on maintaining access to the sources directly.

- **Interpretive context collections** aim to surround a diverse set of exemplary (not necessarily definitive) sources with interpretive context and make interrelationships between sources and context actionable and usable. Sustainability and preservation efforts for such collections would likely prioritize metadata over sources themselves.
- **Evidential platform collections** are focused on aggregating, deconstructing, and remodelling diverse forms of primary sources for new analytical and interpretive uses, for example by deriving computationally amenable data from primary sources. Sustainability and preservation efforts for such collections would likely prioritize the data along with rigorous documentation of provenance and persistent links to original sources.

Of course, many collections combine aspects of each of these varieties of contribution (and presumably many other varieties). If the aim of sustainability and preservation efforts is to maintain the contributions of digital scholarship, then those efforts must be adaptive to varieties of contribution. The digital humanities community lacks a common vocabulary for discussing different modes of contribution of digital scholarship; thus, the first of the four challenges identified in the interviews is a conceptual challenge. The rest of this section elaborates the four challenges identified above.

(1) Discontinuity between the essential interactivity of digital collections and the paradigm of artifactual preservation. Thematic research collections tend to be *essentially* interactive. User-interactivity, collection performativity, or experientiality are often integral to the purposes and intellectual contribution of the collection. Customized browsing functions that exploit scholarly encodings, indexing and navigational schemes that manifest scholarly interpretation, specialized reading and annotation tools, games, interactive maps, three-dimensional models, and simulations – the interactive components of digital collections are often designed to accomplish multiple things at once: to manifest interpretive stances, to enable knowledge transfer, and simultaneously to serve as platforms for ongoing research (Palmer et al., 2009; Fenlon, 2017). Therefore, many collections must remain interactive for their contributions to be manifest. Collections are intended to be “living” (Participant 7). For many collections to be realizing their scholarly purposes, they may not be decomposed into “items,” “objects,” or “raw data,” or reconstructed in a standard content management system.

The interactivity of digital scholarship challenges the prevailing paradigm of artifact-oriented digital preservation. A scholar-centric paradigm of sustainment would prioritize the sustainment of *contributions*, which may be amorphous, and which may or may not neatly align with preservation-ready outputs. There are some promising solutions to aspects of this problem emergent from software preservation and web archiving research (e.g., Rhizome’s Webrecorder¹), which begin to confound the distinction between sustainment and preservation. Indeed, interview participants in this study tended to conflate the terms *sustainability* and *preservation* in light of the essential interactivity of digital collections. One implication of this challenge is the need for a stronger vocabulary for articulating the contributions of digital scholarship to support determinations about what needs to be kept “alive” (and in what form, and for how long), and what can be effectively fixed in amber. It also seems likely that *sustainability* itself will mean very different things to different research communities in different contexts, and this needs further research.

(2) The importance and vulnerability of “connective tissue” within and between collections. Digital humanities collections are networked resources with visible and invisible dependencies among components, and with external resources and services. A collection’s contents may be less essential than “connective tissue” among contents (Participant 5). Connective tissue – interrelationships among components and contextual information, often forged through links or calls to external resources and customized schemas and utilities – may constitute the main interpretive or intellectual contributions of a collection, transcending the discrete digital objects that are the ‘items’ of a collection. However, the same connective tissue is

¹ Rhizome’s Webrecorder: <https://webrecorder.io/>

highly vulnerable to dissolution precisely because it tends to be invisible, undocumented, or technically bespoke and difficult to migrate. This poses the most immediate technical challenge for both sustainability and preservation.

Integral and interstitial components of collections frequently carry important and inexplicit meaning and context. The term *relationships* is used here to indicate constitutive pieces of collections that are not readily classed as primary or secondary sources or data, including links or calls to external data sources and services; implicit contextual and relational information asserted via narrative and design elements; descriptive and relational schemas and ontologies; and computed components such as information retrieval components, dynamic components, algorithmic components, etc. Fenlon (2019) identified a distinction between *direct* and *indirect* relationships undergirding digital scholarship. Direct relationships are referential relationships that are formalized and actionable, for example as calls to URIs encoded in processing scripts or in files, which serve to interrelate, for example, page images to corresponding encoded transcriptions and relevant external standards and authorities. Indirect relationships, on the other hand, are visible and usable in the design of a collection or its web presence (for example, when a webpage juxtaposes a manuscript image with a transcription of the image), but are technically performed by completely unrelated, often computational processes, and are not encoded explicitly in the digital objects comprising the collection. Relationships that are inexplicit or forged dynamically through computation are vulnerable to loss during migration and preservation actions, during staffing changes, and in the absence of thorough documentation. Indirect relationships within a collection's architecture may prove essential to the meaning and the contribution of a collection, and they are intuitively more difficult to characterize and document, let alone sustain or preserve.

One participant, describing how important semantic and editorial information was located in stylesheets rather than directly in digital objects, noted that, “if those things ever get separated, you’ve lost a huge analytical contribution,” and acknowledged the “tight interconnectedness, the integration of purposes of these two things – the phenomena of the data model and the other, related phenomena of the stylesheet or the computational processes” (Participant 8). Becker (2018) has detailed the metaphorical and computational nature of digital objects, and the challenges for preservation work. These challenges are amplified when we consider not only aggregated and interrelated objects, often rife with external dependencies, but also objects that are essentially interactive. This challenge seems likely to grow in an era of linked data and increasingly networked digital scholarship.

(3) Ambiguity of institutional contexts and roles. While many digital humanities collections are created, managed, and sustained by communities of use, they may bear a great variety of relationships to institutional libraries, ranging from complete independence to active and formalized partnership. A collection's institutional context, including factors such as its administrative home within the organization of a university or its proximity to the library, bears heavily on its sustainability, particularly affecting how collection curators are able to plan for or implement maintenance as opposed to innovation or development. This study found that the roles of various entities with a stake in digital scholarship – including scholars, academic departments, libraries and units within libraries, and digital humanities centres and labs – are complex, context-dependent, and subject to ongoing negotiation. Roles within the system of scholarly communication at large become systematized and institutionalized only around established, well understood genres, which may help explain why comparatively unfamiliar or nascent forms of digital scholarship have struggled to attain systematic treatment in libraries.

Participants were unanimous that libraries have a significant role to play in the sustainment of digital scholarship. Most participants reported having had one or more interactions with the library toward the maintenance or preservation of digital humanities collections. Two participants reported that their respective centres had established relationships and standing agreements with the library, which ensure that the library would serve as the “eternal resting place” (Participant 4) for each digital humanities centre's collections, but in both cases the commitment did not carry a timetable for transfer of responsibility, and was constrained to item-level metadata and limited types of items that would fit readily into the existing institutional

repository. Determining transfers of responsibility can be a fraught exercise: it is rarely clear when digital projects are “done and ready for the library to migrate and preserve, and sort of embalm, or whether they were things that the scholar might still like to add to” (Participant 7). Another participant, working within a digital humanities centre, noted that when a centre is physically or administratively located within a library there seems to be an almost unconscious reliance on the surrounding infrastructure to bear the weight of stewardship of collections: “I don’t have to constantly worry about [preservation] because there’s an infrastructure around me that’s thinking about this” (Participant 3). However, no participants reported having established systematic measures or ongoing processes for collaborating with libraries in sustainability and preservation. In some cases where librarians play active roles in the development and maintenance of digital scholarship, their involvement may not reflect established or sustained administrative or institutional support from the library; it may just reflect the initiative of individual librarians. One participant noted that librarians often enter into digital-scholarship collaborations almost “in spite of or around the edges of their existing roles” (Participant 6). Another suggested that digital humanities centres can serve as a “focal point for collaboration between librarians and faculty” toward increasing the library’s roles as “a partner in the research enterprise” (Participant 9).

While libraries continue to increase support for digital scholarship and digital publishing, and indeed take increasingly active roles in research and the collaborative construction of thematic research collections and other forms of digital scholarship, it is not always clear how library digital scholarship initiatives are related to collection development and preservation missions of the library. The appropriate and sustainable division of labour for digital collections is of course a heavily context-dependent determination, and one that may be negotiated and renegotiated over time. As mentioned above, there is no consensus around the value of digital scholarship from an institutional perspective, nor a strong understanding of how libraries or preservation institutions should negotiate the ownership of collaborative and distributed projects (Maron and Pickle, 2014). This study evinces the need for increased research into context-dependent sustainability strategies, and the many and varying roles to be played by different stakeholders.

(4) Lack of infrastructure for collaborative humanities workflows. Emergent digital humanities preservation and sustainability strategies are increasingly prospective. Libraries seek to make interventions earlier in scholars’ development processes, to help scholars make more sustainable technological and representational choices, and to gather requirements to make sustainability plans. As an alternative to the pattern of retrospectively migrating digital projects into the care of libraries after their development, there are increasing efforts to develop and implement common preservation-oriented infrastructures that have the flexibility and extensibility to undergird distributed, custom development by individual digital humanities projects. Prospective strategies aim to lay sustainable foundations in the form of preservation-oriented data management systems underlying advanced indexing and access layers, as platforms on top of which humanists can build expressive, interpretive, customized digital scholarship (e.g., Sweeney et al., 2017; Madsen and Hurst, 2018; Almas, 2017; White et al., 2019).

The success of cyberinfrastructure for the humanities will depend on its capacity to accommodate the wide-ranging human and technical processes or *workflows* that structure the development and maintenance of collections. Indeed, we can understand those workflows as integral to the infrastructure of collections, and therefore of sustainability. The workflows or processes that create and maintain collections (and digital humanities scholarship generally) are idiosyncratic, distributed, and highly collaborative, and this will complicate attempts to establish a shared cyberinfrastructure even within domains of research (Fenlon, 2019). Indeed, this study found that beyond maintenance of the technical components of a collection, sustaining a collection may depend on the maintenance of human workflows. One interview participant described needing to alter the course of a whole collection-development workflow – a distributed and collaborative process of digitization, transcription, and ingest – in order to conduct a routine data migration. This participant described the difficulty and necessity of implementing

changes to a workflow that was well established and distributed across teams at multiple institutions, asserting that alterations to workflow necessarily accompany technical maintenance and may in fact be more complex: “having a conversation about...what the folks working on [the collection] *like* to do, want to do with it—that *was* sustainability work – and keeping their workflow intact in some ways, but just fixing some things that maybe weren’t working” (Participant 1).

Toward Community-Centred Sustainability Strategies

This study illuminated several institutional strategies for digital humanities scholarship, some of which are well established in library practice, while others are emergent. As described above, the most common, institutionally centred strategies for sustainability and preservation rely on negotiated levels of commitment and, ultimately, handoff of responsibility for the collection from the original creators to a curation institution, often with some loss of fidelity to the collection. This strategy is inevitably inadequate for handling the diversity and scope of digital scholarship, due to the challenges described above: comprehensive collection of digital scholarship would exceed the capacity of most preservation institutions; and there are aspects of digital scholarship that strongly resist common approaches to preservation or shared, scalable curatorial and research infrastructures.

This research identified a promising complement or alternative to institutionalized sustainability strategies: reorienting sustainability efforts toward research communities, rather than focusing exclusively on collections themselves. The notion of community-centred sustainability emerges from two interrelated outcomes of the interviews: (1) collection sustainability depends on engaging communities of interest, including original creator and user communities, development/maintenance communities, and communities of reuse; and (2) as described above, maintaining collections may frequently entail maintaining the sociotechnical workflows that structure collaborations within research and development communities.

Interview participants were unanimous about the critical importance of *use* to ensuring a collection’s sustainability. One participant observed that stakeholder engagement is more important than any technical intervention: “the bigger concern is not, How do you structure these?...It’s really, How do you create those kinds of community engagements that result in people squawking if the project goes away?” (Participant 7). The study suggested strong interest among collection stakeholders in the strategy of preparing collections to pivot toward new purposes and therefore new user communities over time. One participant suggested that collections might be documented and structured from the start to support handoffs to new research communities, mirroring patterns of open-source software development. However, this participant also acknowledged significant obstacles, including the lack of support and incentive in digital humanities research for repurposing existing collections rather than developing new ones (Participant 1). Participants also suggested that aggregating thematically related collections might help combine and grow user communities from across disciplines or topical areas.

Community-centred sustainability strategies revolve around the ongoing growth and development of collections in service to communities, further highlighting the distinction between *sustainability* and *preservation* of digital humanities scholarship. The idea that purposefully and strategically growing and engaging user communities benefits the sustainability of collections is not new. In a study of open data and digital curation practices, Lee et al. (2016) argued that the mission of the curator must be extended beyond access-provision to the facilitation of new forms of use and interaction with and among users of data. In addition, Post (2017) has explored new models of institutional and community partnership for the preservation of new media art. However, the question of how curators can purposefully grow community engagement with a collection or, alternatively, increase the capacity of collections for use and development by varying communities, remains open and vitally important to the future of

humanities data curation. Despite a robust literature on humanities scholars' information practices, ongoing digital curation efforts would benefit from increased understanding of the needs of users of digital humanities scholarship and scholar-generated collections specifically.

Future Work

By re-orienting our conception of sustainability toward research communities rather than focusing exclusively on the collections or artifacts created and used by those communities, we open a landscape of possibilities for collaborative sustainment of digital scholarship. Community-centred archiving strategies, including community-oriented acquisition and participatory archives, aim to reorient archival practice away from institutional imperatives and toward the well-being and endurance of communities (Christen and Anderson, 2019; Caswell and Cifor, 2016; Gilliland and Flinn, 2013; Caswell, 2014; Yoon, 2013; Shilton and Srinivasan, 2007). In cultural heritage practice, there are numerous emerging models of institutional partnership with communities, including efforts to:

- Create resources such as toolkits, workshops, and community-oriented best practices to support community curation work;
- Provision community sustainability efforts through re-granting programs, the reallocation of collection development funds toward community investments, or in-kind resources such as library staff time and consultation;
- Establish spaces and practices for building trust and equitable partnership among communities and memory institutions;
- Develop a common foundation of principles along with model policies and agreements toward ongoing partnership or shared stewardship.

While many of these developments are happening in the context of community archives theory and practice in cultural institutions, rather than in the realm of digital scholarship and academic libraries, there is significant commonality across community archives and digital humanities collections (centred in research communities), and in the sustainability challenges they face. Future work will explore the overlap among and differences between collections centred in different kinds of communities, and the sustainability strategies available to them.

The results reported here have laid the groundwork for an ongoing investigation into the sustainability challenges confronting collections more broadly, particularly collections that are created, managed, and sustained primarily by their communities of use, either outside of the purview of memory institutions or in tentative or provisional relationships with memory institutions. Future work aims to support and extend this movement toward community-centred sustainability of all kinds of digital collections through case studies of digital humanities collaborations and collections. The goal of future work is to answer foundational questions confronting next-generation sociotechnical infrastructures for long-lived cultural and scholarly records: on what *sustainability* means for different communities, different stakeholders within communities, and different collection contexts; on the contributions, purposes, and *completeness* of different forms of digital scholarship; and around the distinctive and evolving roles of institutions and communities in sustaining cultural records.

References

- Almas, B. (2017). Perseids: Experimenting with infrastructure for creating and sharing research data in the digital humanities. *Data Science Journal*, 16(0). doi:10.5334/dsj-2017-019
- Caswell, M. (2014). Community-centred collecting: Finding out what communities want from community archives. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–9. doi:10.1002/meet.2014.14505101027
- Caswell, M., & Cifor, M. (2016). From human rights to feminist ethics: Radical empathy in the archives. *Archivaria*, 81. Retrieved from <https://archivaria.ca/index.php/archivaria/article/view/13557>
- Christen, K., & Anderson, J. (2019). Toward slow archives. *Archival Science*, 19(2), 87–116. doi:10.1007/s10502-019-09307-x
- Clement, T., Hagenmaier, W., & Knies, J. L. (2013). Toward a notion of the archive of the future: Impressions of practice by librarians, archivists, and digital humanities scholars. *The Library Quarterly*, 83(2), 112–130. doi:10.1086/669550
- Cooper, D., & Rieger, O. Y. (2018). *Scholars ARE collectors: A proposal for re-thinking research support* [Issue brief]. Retrieved from Ithaka S+R website: doi:10.18665/sr.310702
- Dombrowski, Q. (2014). What ever happened to Project Bamboo? *Literary and Linguistic Computing*, 29(3), 326–339. doi:10.1093/lc/fqu026
- Eschenfelder, K. R., Shankar, K., Williams, R., Lanham, A., Salo, D., & Zhang, M. (2016). What are we talking about when we talk about sustainability of digital archives, repositories and libraries? *Proceedings of the 79th ASIS&T Annual Meeting: Creating Knowledge, Enhancing Lives Through Information & Technology*, 148:1–148:6. Retrieved from <http://dl.acm.org/citation.cfm?id=3017447.3017595>
- Fenlon, K. (2017). *Thematic research collections: Libraries and the evolution of alternative scholarly publishing in the humanities* (Doctoral dissertation, University of Illinois at Urbana-Champaign). Retrieved from <http://hdl.handle.net/2142/99380>
- Fenlon, K. (2019). Modeling digital humanities collections as research objects. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, Urbana-Champaign.
- Flanders, J. (2014). Rethinking collections. In P. L. Arthur & K. Bode (Eds.), *Advancing Digital Humanities* (pp. 163–174). doi:10.1057/9781137337016_11
- Gilliland, A., & Flinn, A. (2013). *Community archives: What are we really talking about?* Presented at the CIRN Prato Community Informatics Conference 2013. Retrieved from https://www.monash.edu/__data/assets/pdf_file/0007/920626/gilliland_flinn_keynote.pdf
- Goddard, L., & Walde, C. (2017). Negotiating sustainability: The grant services “menu” at UVic Libraries. Presented at the Digital Humanities (DH2017). Retrieved from <https://dh2017.adho.org/abstracts/231/231.pdf>

- Green, H. E., & Courtney, A. (2015). Beyond the scanned image: A needs assessment of scholarly users of digital collections. *College and Research Libraries*, 76(5), 690-707. Retrieved from <http://crl.acrl.org/content/early/2014/09/08/crl14-612.full.pdf>
- Langmead, A., Quigley, A., Gunn, C., Hakimi, J., & Decker, L. (2018). *Sustaining MedArt: The impact of socio-technical factors on digital preservation strategies* (Report of Research Funded by the National Endowment for the Humanities, Division of Preservation and Access No. PR-234292-16). Retrieved from https://sites.haa.pitt.edu/sustainabilityroadmap/wp-content/uploads/sites/10/2017/01/SustainingMedArt_FinalReport_Web.pdf
- Lee, C., Allard, S., McGovern, N., & Bishop, A. (2016). Open data meets digital curation: An investigation of practices and needs. *International Journal of Digital Curation*, 11(2). doi:10.2218/ijdc.v11i2.403
- Madsen, C., & Hurst, M. (2018). Are digital humanities projects sustainable? A proposed service model for a DH infrastructure. Presented at the Coalition for Networked Information Fall Membership Meeting (CNI18f), Washington, D.C. Retrieved from <https://www.slideshare.net/mccarthymadsen/are-digital-humanities-projects-sustainable-a-proposed-service-model-for-a-dh-infrastructure>
- Maron, N. L., & Pickle, S. (2014). *Sustaining the digital humanities: Host institution support beyond the start-up phase* [Research Report]. Retrieved from ITHAKA S+R website: <https://sr.ithaka.org/publications/sustaining-the-digital-humanities/>
- Oltmanns, E., Hasler, T., Peters-Kottig, W., & Kuper, H.-G. (2019). Different preservation levels: The case of scholarly digital editions. *Data Science Journal*, 18(1), 51. doi:10.5334/dsj-2019-051
- Palmer, C. (2004). Thematic research collections. In *A Companion to Digital Humanities*. Blackwell Publishing. Retrieved from http://digitalhumanities.org:3030/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-5&toc.depth=1&toc.id=ss1-4-5&brand=9781405103213_brand
- Palmer, C. L., Tefteau, L. C., & Pirmann, C. M. (2009). *Scholarly information practices in the online environment: Themes from the literature and implications for library service development*. Retrieved from OCLC Research and Programs website: <http://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf>
- Post, C. (2017). Preservation practices of new media artists: Challenges, strategies, and attitudes in the personal management of artworks. *Journal of Documentation*, 73(4), 716-732. doi:10.1108/JD-09-2016-0116
- Prescott, A. (2015). Beyond the Digital Humanities Center. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A New Companion to Digital Humanities* (pp. 459-475). Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/9781118680605.ch32/summary>
- Rawson, K., & Muñoz, T. (2019). Against cleaning. In M. K. Gold & L. F. Klein (Eds.), *Debates in the Digital Humanities*. Retrieved from <https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/07154de9-4903-428e-9c61-7a92a6f22e51>
- Shilton, K., & Srinivasan, R. (2007). Participatory appraisal and arrangement for multicultural archival collections. *Archivaria*, 63(0), 87-101.

- Smithies, J., Westling, C., Sichani, A.-M., Mellen, P., & Ciula, A. (2019). Managing 100 digital humanities projects: Digital scholarship & archiving in King's Digital Lab. *Digital Humanities Quarterly*, 13(1).
- Smithsonian Center for Folklife and Cultural Heritage. (2019). Shared stewardship of collections policy. Retrieved October 22, 2019, from <https://folklife-media.si.edu/docs/folklife/Shared-Stewardship.pdf>
- Sustaining Digital Scholarship. (2004). *SDS Final Report*. Retrieved from Institute for Advanced Technology in the Humanities website: https://dcs.library.virginia.edu/files/2012/05/SDS_FinalReport2003.pdf
- Sweeney, S. J., Flanders, J., & Levesque, A. (2017). Community-Enhanced Repository for Engaged Scholarship: A case study on supporting digital humanities research. *College & Undergraduate Libraries*, 24(2-4), 322-336. doi:10.1080/10691316.2017.1336144
- Vinopal, J., & McCormick, M. (2013). Supporting digital scholarship in research libraries: Scalability and sustainability. *Journal of Library Administration*, 53(1), 27-42. doi:10.1080/01930826.2013.756689
- White, L., Feldman, K., Kosnik, A. D., Gere, A., Naglak, M., Rathemacher, A., & Cohen, S. (2019, October). *Sustainable publishing for digital scholarship in the humanities*. Technical Services Faculty Presentation presented at the Charleston Conference Webinar Series. Retrieved from https://digitalcommons.uri.edu/lib_ts_presentations/71
- Yoon, A. (2013). Defining what matters when preserving web-based personal digital collections: Listening to bloggers. *International Journal of Digital Curation*, 8(1). Retrieved from <http://www.ijdc.net/article/view/8.1.173>