

Sustaining Digital Community Collections:
Understanding the Impact of Development Workflows Across Contexts

Digital collections created by humanities scholars and community archives are extensive, scattered bodies of evidence of cultures and histories that are often underrepresented in cultural institutions. Born outside of memory institutions, these collections confront major barriers to sustainability, compromising the completeness and equity of our collective digital memory. “Sustaining Digital Community Collections” is a three-year, \$380,646 Early Career Development research project (October 2020–September 2023) led by Dr. Katrina Fenlon (University of Maryland College of Information Studies) to investigate new models for how memory institutions may support the sustainability of community digital collections. Libraries and archives cannot and should not comprehensively collect and sustain the growing mass of digital community collections—particularly where communities wish to retain some level of power over the ongoing care and use of collections. Developing context-driven sustainability models relies on identifying and assessing the human and technological workflows that create and maintain collections over time. This project investigates collaborative workflows of collection development and maintenance, to identify roles that cultural heritage institutions may play to help realize community-determined, community-led strategies for sustaining digital collections. This project will develop and evaluate: (1) Models of emergent development and maintenance workflows for digital humanities research collections and digital community archives; (2) Roles for libraries and archives in those workflows; and (3) A toolkit to help stakeholder communities and memory institutions map sustainable, equitable workflows and identify context-driven, community-centered strategies for collection sustainability.

Statement of broad need. The sustainability of digital cultural collections is an urgent problem that transcends collecting institutions and communities of research and memory. There are hundreds or perhaps thousands of digital collecting projects in active development among scholarly and public communities, many of which provide access to unique evidence (data and digital primary sources) of underrepresented people and cultures. For example, the [Enslaved project](#) is building a linked open data platform for the study of the historic slave trade; the [Open Islamicate Texts Initiative](#) is building a corpus for the study of texts from the premodern Islamicate world; and the [Lakeland Digital Archive](#) is a partnership between the Lakeland Community Heritage Project and the Maryland Institute for Technology in the Humanities to document a local, historic African American community. Most community collections are developed independently of preservation institutions, funded sporadically on short cycles, and rely on bespoke technical infrastructures. These factors combine to make sustainability uniquely problematic for community-centered digital collections. Though many at-risk collections are within scope of preservation missions of institutional libraries and archives, most institutions lack capacity to take in and sustain any more than a narrow swath of community collections. Beyond that, communities have deep investments in their collections; many seek new models of institutional partnership that keep varying levels of power over and responsibility for collections in the community’s hands. The need for sustainable infrastructures built on values of openness, equity, and partnership has been acknowledged as a grand challenge for knowledge ecosystems of the future.¹ This project builds on a prior study of digital humanities collections, which found that sustaining digital collections depends on understanding and maintaining idiosyncratic, distributed, collaborative workflows of collection development and maintenance. In order to build sustainable infrastructures for the cultural record in a growing diversity of communities and institutions, we need an improved understanding of the human and technical workflows on which collections depend.

Project design. This three-year project seeks to answer the following questions: (1) How can we identify, characterize, and represent emerging sociotechnical workflows for the development and ongoing maintenance of digital collections, including digital humanities collections and community archives? (2) How do these workflows affect the sustainability of collections in different contexts? (3) What are roles and best practices for

¹ <https://grandchallenges.pubpub.org/pub/final>

libraries, archives, and stakeholder communities in supporting community-centered sustainability workflows for collections? This project will address these questions through a comparative, multi-case study of three central cases (all described above): (a) the *Enslaved project*; (b) the *Open Islamicate Texts Initiative*; and (c) the *Lakeland Digital Archive*. Each case is currently developing a collection about underrepresented people, literatures, or communities, and each case bears a different relationship to relevant libraries and archives. The PI has an established research relationship with each community; all have given assurances of support, access, and partnership in this work. Specific activities are as follows:

Year 1. Develop case studies of community collection workflows in active development. The project team will develop case studies of each collection in development, with a focus on characterizing sociotechnical, collaborative workflows of development and planned maintenance, including for metadata-creation, data migration and transformation, and provenance documentation. Case studies will be developed through interviews, document analysis of planning and technical documentation, workflow mapping, and site visits to conduct participant-observation in development meetings. The project team will draft initial models of workflows, identifying points of vulnerability for sustainability. **Year 2. Build cross-case comparisons with established community collections and memory institutions.** The project team will augment the three central cases with interviews with representatives of relevant university libraries and archives that have potential stakes in each collection, to identify roles for library and archives professionals. In addition, through semi-structured interviews and analysis of documentation, the project will build 5-10 additional mini-cases of established, long-running digital humanities collections or digital community archives, to complement the cases of in-development collections through retrospective analysis of successful sustainability efforts. **Year 3. Formalize and evaluate workflows and toolkit for communities and libraries.** Through cross-case analysis, the team will formally describe workflows for sustainable collection development and maintenance, and formally evaluate workflows in partnership with the case study communities. The project team will create and publish an open toolkit for stakeholder communities and memory institutions to map equitable workflows and context-responsive, community-centered strategies for sustainable collection development and maintenance.

Diversity plan. Because the driving impetus for this project is to sustain collections within a growing diversity of communities and institutions, recruiting researchers from underrepresented populations in information professions is an express priority. The University of Maryland iSchool is an established leader in diversity and inclusion. Student researchers will be recruited from across iSchool programs, including our unprecedented digital curation specialization for undergraduates, which has a highly diverse student population, and which aims to open pathways for diverse students into library and archive professions.

Broad impact. The project will generate insights about digital collection sustainability that transcend institutional contexts, to help increase the sustainability of digital scholarship in the humanities and digital community archives, and to provide memory institutions with navigable paths forward in supporting the viability of diverse, distributed digital collections. The proposed work will contribute to theory and practice in digital curation, digital humanities, scholarly communication, and knowledge organization and representation.

Budget. The proposed budget of \$380,646 entails \$162,996 for student support (graduate researcher tuition support, stipends, and benefits, and undergraduate hourly wages); \$15,000 for travel, to support PI and student site visits for data collection, and conference travel; \$4,500 for materials and supplies; \$78,654 for 3 years of PI summer salary and fringe benefits; and \$119,496 in indirect costs at the negotiated on-campus rate of 54.5%.

PI. Dr. Fenlon is an assistant professor at the College of Information Studies at the University of Maryland, College Park, and an affiliate of the Maryland Institute for Technology in the Humanities. Her research focuses on the curation, representation, use, and sustainability digital cultural collections. She employs qualitative and information modeling methods to study the social and technical challenges of sustainable, shared infrastructure for the humanities. Her recent work investigates data models at the intersection of linked cultural data and humanities scholarly communication.