

Abstract

The Northeastern University Library's Digital Scholarship Group seeks funding for a national forum to inform and support the development of a teaching and learning toolkit focused on a set of questions concerning the handling of diversity within the design of information management systems, interfaces, and protocols. How do knowledge representation, naming, and power influence not only interactions with disenfranchised communities, but also the design of tools and interfaces? When tools and interfaces limit or constrain interactions with documents and items, are they responsive to community needs or do they truncate ways of knowing? What commonalities exist (or turn out not to exist) across communities of users? Questions like these will be discussed by participants drawn from a broad range of communities at an opening forum event. The results will be synthesized and analyzed at a second meeting by a small design group who will develop a draft toolkit of teaching and learning materials. Those materials will be circulated for public input and comment, and discussed in detail at a concluding open forum, whose results will be used to create a final version of the toolkit. The project is planned as a two-year initiative running from December 2016 through November 2018.

This project addresses important needs within both the libraries, archives, and museum (LAM) community and the many communities whose cultural heritage is at stake. The questions we seek to address are inherent in the turn to the digital and exacerbated by the renewed focus on community archives, both within the LAM world and without. Increasingly, researchers across all disciplines seek to capture and preserve the voices of communities not often given a voice, communities which have often experienced forms of trauma. Yet the tools of preservation and dissemination draw strongly on traditions and systems of information management that arise from colonizing and homogenizing imperatives. Pragmatic considerations such as broad accessibility and discoverability may operate in tension with the need to accommodate local knowledge, non-hegemonic modes of information organization, and culturally sensitive handling of specific materials, even in cases where those pragmatic concerns arise from within the communities affected. The challenge is not to simply overturn established knowledge management systems, but to find modes of critique and accommodation that enable those systems to be used by a much broader constituency of users, without entailing the erasure or misrepresentation of cultural meaning.

The audience for this project is three-fold: first, those within the LAM community who are currently engaged in grappling with this set of issues; second, information and cultural heritage curators in all contexts and at any level for whom the toolkit may prove a valuable source of impetus for teaching and learning; and third, those in culturally disenfranchised communities whose relationship with and access to information systems we hope, over the long term, to improve. We intend the toolkit, and the results of the forum discussions, as a teaching and learning resource that will inform the next generation of development, resulting in systems that are less rigid and more structurally accommodating of diverse forms of knowledge.

Narrative

Statement of Need

The acquisition, housing, and care of cultural objects has always had colonial implications. The library, archives, and museum (LAM) world has begun to grapple more openly with the consequences of asserting control over items removed, often forcibly, from their original cultural context. Professional LAM communities now assert the desire for genuine, responsive partnership with the communities in which these artifacts were created¹, deeper understanding of the hegemonic role of knowledge representation via standardized ontological decisions², and investigation into the role of algorithm, interface, and tool design in reinforcing power differentials inherent in the status quo.³

These questions are inherent in the turn to the digital and exacerbated by the renewed focus on community archives, both within the LAM world and without. Increasingly, researchers across all disciplines seek to capture and preserve the voices of communities not often given a voice, communities which have often experienced forms of trauma. Now that academic researchers create and theorize “the archive” at an ever-increasing pace, what are the responsibilities of the researcher, the designer, the curator, or the archivist in creating new open digital collections ready for automated harvesting, software-assisted analysis, and broad re-use and reinterpretation? How do knowledge representation, naming, and power influence not only interactions with disenfranchised communities, but also the design of tools and interfaces? When tools and interfaces limit or constrain interactions with documents and items, are they responsive to community needs or do they truncate ways of knowing? What commonalities and diversities exist within and across communities of users?

We seek funding for a National Forum to begin to answer those questions. The forum events will bring together not only librarians, archivists, and curators, but also historians and oral historians, humanists and digital humanists, programmers and metadata experts, anthropologists and sociologists -- all people with a vested interest in the ethical creation and support of digital archives involving disenfranchised or underrepresented communities⁴. Most importantly, these events will include members of disenfranchised communities who are seeking themselves to create and manage their own digital archives, and to intervene in and contribute to archives maintained by others.

Community engagement, recovering history, and collecting and archiving the voices of disenfranchised and underrepresented groups are all topics of the moment, both within and outside the LAM community. Important and relevant work has already been supported by the

¹ See, for example, the repatriation policies of the United States’ National Museum of the American Indian or Johnston on recent discussions of repatriation prompted by new publications in the museum field; full citations for all references available in our **References and Works Consulted** in the supporting documents.

² See, for example, Berman and Olson.

³ See, for example, McPherson, Sweeney, and Chun.

⁴ A note about language: we use the terms “disenfranchised”, “marginalized”, or “underrepresented” to describe groups that have been actively excluded from the priorities of predominantly white institutions. We recognize that this carries the risk of further distancing our work from those communities, and places historically white institutions as central actors. However, in LAM and related fields, “marginalized” and “underrepresented” are terms of art used to specifically point towards communities damaged by the destructive mechanisms of systemic bias, and we wish to acknowledge the active role historically white institutions have had in disenfranchising these groups.

IMLS⁵, and we can see the importance of these topics not only in a scan of academic work⁶ but also in recent venues across disciplines such as journal special issues⁷, conference presentations and keynotes⁸, and current CFPs⁹. We also see new projects, again headed by interdisciplinary groups, focused on building new methods, workflows, and tools for archiving new and born-digital web and social media content specifically as a corrective to past silences in the archive.¹⁰

These efforts have produced valuable outcomes in the form of best practices documentation, expert knowledge in specific areas, important and inclusive digital archives, professional development within the LAM community, and effective outreach and education efforts targeted at specific groups. However, guidance on best practices does not always reach a cross-disciplinary audience, and has also not always proved effective as a way of reaching those outside the direct expertise community of those actively studying the problem. To extend the reach of this work, we will first convene an Opening Forum that draws on the broad base of existing work and brings in discussants from varying domains to identify remaining areas of challenge. We will then synthesize the results of that discussion into a teaching and learning toolkit that can be used in a variety of contexts, and present the results of that work for discussion and formative critique at a concluding Review Forum. We envision the toolkit as an adaptable, cross-disciplinary resource consisting of small-scale units (case studies, readings, individual assignments, discussion questions) plus a set of higher-level units to organize those smaller units into narratives (e.g. a two-day workshop, a three-week course module, etc.) While developing a full-fledged, programmatic curriculum is outside the scope of a National Forum, we expect forum activities and the resulting materials and framework to prompt further development within the relevant educational fields. Finally, this initiative also seeks a more pointed outcome: to serve as a corrective to neoliberal impulses in both DH and LAM in which simplistic forms of technological and economic development are construed as self-evident goals. The discourse of “libraries of the future” risks echoing elements of Allington et al.’s admittedly broad-brush and pessimistic characterization of digital humanities as a field where “discourse sees technological innovation as an end in itself and equates the development of disruptive business models with political progress.”¹¹ This forum aims at the critical infrastructure approach suggested by Alan Liu in his discussion of the future direction of digital humanities¹², in which he calls for “shared but contested information-technology infrastructures”: knowledge systems that work as infrastructure but do not suture over cultural difference. While thorough solutions to this problem are out of scope for this forum, we do seek to generate an inclusive framework for *investigating* and

⁵ We are particularly inspired by three recently-funded IMLS projects, the *Digital Atlas of Native American Intellectual Traditions*, the *Tribal Stewardship Cohort Program*, and the upcoming forum “Diversifying the Digital Historical Record”, as well as the structural model of the Sustainable Heritage Network more generally.

⁶ See the **References and Works Consulted** in our Supporting Documents for more detail on this and other citations.

⁷ See, for example: *Code4Lib Special Issue on Diversity in Library Technology*, 2015; *Archival Science Special Issue: Keeping Cultures Alive: Archives and Indigenous Human Rights*, 2012; *Library Quarterly Special Issue on Diversity and Library and Information Science Education*, 2013.

⁸ See Safiya Noble’s *Power, privilege and the imperative to act* keynote at the Digital Library Federation conference (Fall 2015), Mark Matienzo’s *To Hell With Good Intentions: Linked Data, Community and the Power to Name* keynote at the LITA Forum (Fall 2015), Johnetta Betsch Cole’s *Museums, Diversity, and Social Change* at the American Alliance of Museums (Spring 2015) and Society of American Archivists panels and fora like *Mind Your Own F#@king Business”: Documenting Communities that Don’t Want to Be Documented and the Diversity of the American Record* (2015) and *Diversifying the Archival Record* (2014).

⁹ *Digital Humanities Quarterly* CFPs for special issues on Spanish-language and French-language Digital Humanities; *Archives and Records* CFP for a special issue on public history, archives, and identity; ALISE '17 CFP on community engagement and social responsibility in LIS education.

¹⁰ Documenting the Now, Social Media Archives Toolkit, Documenting Ferguson, Our Marathon, to name a few.

¹¹ Allington, D., Brouillette, S., & Golumbia, D. “Neoliberal Tools (and Archives): A Political History of Digital Humanities.”

¹² Liu, Alan. “Drafts for Against the Cultural Singularity (book in progress).”

teaching the problem so that future infrastructural work proceeds in a critical spirit.

This work will be shaped by both local and national expertise. The core of our Advisory Board is drawn from local partners with experience in working with communities that have been disenfranchised by historically white, English-speaking institutions and organizations. Our Board includes members with experience in working with tribal communities to preserve indigenous languages (Dr. Ellen Cushman), ethical research partnerships with marginalized communities (Dr. Moya Bailey), preserving and sharing materials on Boston social justice and community history (Giordana Mecagni), public history approaches to community engagement around culturally sensitive topics and materials (Marty Blatt), and global diversity in digital humanities (Dr. Élika Ortega Guzman). We discuss our Board's qualifications further in the **Project Resources: Personnel** section.

We will broaden this base of expertise further with advisory board members, keynote speakers, and discussion leaders who bring national perspectives in a range of relevant domains, including culturally sensitive design (e.g. the Mukurtu CMS); the creation of copyright and re-use licenses sensitive to the needs of indigenous materials¹³; ontologies and data modeling, including cultural heritage groups within the standards community¹⁴; archives, libraries, and data and other curation at a range of large and small cultural institutions; and community work (outside the traditional LAM environment) on collecting and digitizing at-risk heritage materials. Keynote speakers and discussion leaders whom we seek to recruit include Wendy Chun, Hope Olson, Angel Nieves, Kim Christen Withey, Marisa Elena Duarte, and Kelcy Shepherd; we have already received a supportive response from scholars and practitioners Wendy Chun (*Programmed Visions: Software and Memory*), Angel Nieves (*Apartheid Heritages, Soweto '76*), Kelcy Shepherd (the *Digital Atlas of Native American Intellectual Traditions*), and Kim Christen Withey (*Mukurtu, The Sustainable Heritage Network*, and *The Tribal Stewardship Cohort Program*), and expect to involve them in the project moving forward.

Our audience consists of practitioners, researchers, students, and educators in LAM and related fields. We use the word “practitioners” deliberately, to involve people and organizations involved in LAM work, whether part of a large organization or smaller, grass-roots community efforts. We also define students (in formal LAM programs or not) as including both practitioners and anyone interested in learning more about design for culturally sensitive materials. While focusing primarily on the LAM fields, we also engage with related fields such as digital humanities, public humanities, and digital history.

Impact

The activities of this grant will have impact at several levels. The forum events will convene a high-profile discussion that is open to wide participation (both physical and virtual), raising awareness among those new to the issues and enabling synthesis among strands of existing work. The toolkit will have an important professional development impact, serving as the basis for workshops and helping those who are wrestling with these design challenges in their regular work. It will also have an impact on the next generation of LAM professionals: the toolkit can help bring these issues into the pedagogical context of degree and certificate programs and make them both familiar and urgent. The concluding white paper and the record of the forum events will serve as an important resource for LAM professionals. Finally, we anticipate an important impact on community stakeholders, enabling them to build better connections between their work as curators and designers and the analogous work being done in a more formal LAM context.

¹³ See, for example, prior work on Indigenous Knowledge in New Zealand or Traditional Knowledge labels developed through the Local Contexts program.

¹⁴ CIDOC CRM, ANSI/NISO Adaptive XML.

The focus on a pedagogical outcome is important to this impact. Despite the breadth of previous work in this area, we do not yet see substantial interventions early in the educational or professional development cycle. Issues of social justice, community partnerships, and diverse archives and museums are visible in the formal LAM curricula, but the *information and system design aspects* are not yet a core part of the education of future programmers, catalogers, or interface designers in the LAM world. In a preliminary scan of Master's level library and archives program requirements, we saw that while several institutions have courses that are likely to include material on diversity and inclusion in digital archives, none are required as part of the degree¹⁵, with the notable exception of the Certificate in Community Informatics specialization at the University of Illinois at Urbana-Champaign. Required and elective courses in that certificate likely touch on issues of diversity and inclusion, but we did not see evidence of a focus on information and system design. Master's level museum and museology programs similarly did not seem to have any courses entirely dedicated to the topic, or if they did, those courses were electives.¹⁶ One program in our scan, the MA in Museum Studies at the University of San Francisco, was notable for a required course on culturally sensitive material. However, even there we did not see the emphasis on information and system design that we propose. Significantly, although we did identify several well-established MLS/MLIS programs that cover indigenous knowledge and culture (notably at the University of Washington and the University of Arizona), these programs nonetheless treated that domain as separate from the mainstream of LIS education, handling it as a specialized track rather than bringing this perspective into the required and structurally integral parts of the standard LIS degree program.

We therefore wish to prompt additional work on how to build issues that often underlie social justice, diversity, and inclusion efforts -- at the information and system design layer -- into the set of professional interests and skills at a very early stage. As Kimberly Christen Withey states, when discussing previous efforts of scholars and practitioners working with Indigenous knowledge systems: "Instead of simply "adding" Native comments to the records, or offering an "inclusion" model that only alters Indigenous recordkeeping, [these efforts] suggest structural changes to the core of information management systems across the board."¹⁷

Finally, while this work will be captured in a white paper, we also recognize that the white paper as a genre has a narrow reach. An adaptable pedagogical tool can reach a broader audience than an expert document, and also engages readers as active participants, working through case studies and research questions. A model program in this area is the *Tribal Stewardship Cohort Program*, developed by leads from the Mukurtu project, Washington State University, and the Sustainable Heritage Network, and also funded by the IMLS. Programs like the *Tribal Stewardship Cohort Program* begin with a focus on the collection and management side of archives, and move into reflecting community norms through existing tools. We have reached out to the Mukurtu leads to learn from their experience in both systems and curricular design as well as creating community partnerships and hope to involve them more directly in the forum.

This proposal addresses themes that recur across both IMLS' National Digital Platform and Learning in Libraries agency priorities: inclusive, radical collaborations to increase diversity and community engagement. All three reports from the 2015 IMLS Focus Conference (Engaging Communities, Learning in Libraries, National Digital Platform) emphasize partnership with local

¹⁵ Specific courses on Archives and Cultural Outreach were available electives at Simmons and UNC-Chapel Hill, but we did not see readily apparent course content at other top library science and archives programs like the University of Illinois Urbana-Champaign and University of Michigan Ann Arbor. The University of Texas at Austin and University of Washington programs, which both require a course on Information in Social and Cultural Context and Information and Society respectively, are exceptions in that those courses likely touch on issues of diversity and inclusion, though again are not focused on system design.

¹⁶ For example, those at Syracuse University, New York University, and the University of Washington.

¹⁷ Christen Withey, Kimberly. "Tribal Archives, Traditional Knowledge, and Local Contexts: Why the "s" Matters"

communities in support of participatory learning, digital literacy and broad inclusion. This proposal, focusing on best practices in systems and information design, is most directly relevant to the National Digital Platform priority. By concentrating on increasing education around the elements of system design, this forum complements work already underway on diversifying front-end content, collection, and partnership methods. This forum series and outcomes will therefore serve several key themes that emerged from the National Digital Platform IMLS Focus Report: be inclusive, initiate radical collaborations, and shift to continuous professional learning.

Project Design

Objectives: The principal objective for this project is to convene a national discussion concerning the ways in which information design and management are shaped by forms of cultural hegemony, how that shaping affects the capabilities of information systems to represent and engage a diversity of cultural materials and agents, and how the LAM community might respond with alternative design and usage strategies. The project also seeks to build the results of that discussion into a teaching and learning toolkit to ensure their widest possible impact.

Participants: This forum allows for several types of participation. The two public forum events (Opening Forum and concluding Review Forum) will be free and open to any attendee. We will record and, if possible, stream all public forum events, and encourage open synchronous and asynchronous participation via methods such as Twitter, shared collaborative documents, and surveys or other feedback mechanisms. We will provide generous travel subsidies to up to 25 participants at the Opening Forum, to reduce the effect of inequities in either individual or organizational financial resources. While some participants will be invited, we also plan to widely publicize participation opportunities and encourage potential participants to apply for travel subsidies. Please see the **Diversity Plan** for more detail on recruitment.

We have laid the groundwork for an **Advisory Board (6-10 members, unpaid)** and **Core Design Group (10 members, with honorarium)** who may attend and give presentations at the opening and closing forums and contribute expertise to the design and drafting of the toolkit. Members of the AB may also serve as CDG members if they wish to be more directly involved in drafting the toolkit and attending the toolkit development meeting. We will engage with members of both groups using virtual methods such as email, Skype, and shared writing tools. **The Core Design Group** will carry out the substantive work of actually developing the teaching toolkit, with support from the Graduate Research Assistant, co-PIs, and Repository Manager.

Event Design: We plan three events: the Opening Forum, the Core Design Group Working Meeting, and the Concluding Review Forum. (Sample schedules for these events are included in the supporting materials.) The **Opening Forum** has several goals. It will be designed around principles of inclusion and diversity, and serve as a large watershed for theory, ideas, and sources. (Participant recruitment and design for inclusion is addressed in the Diversity Plan.) We will strike a balance between presentations from experts and provocative voices and opportunities for discussion in small and large groups. On a more practical level, it will serve as a recruitment event for participants in later activities, particularly members of the Core Design Group. The Opening Forum will be organized around three topics: language and problems of naming (including nomenclature, controlled descriptions, cultural specificity, colonial and other hegemonic control through naming); data design and management (data models and information and system design, data as it lives behind the scenes); and interface design and user interaction (the interface as it shapes and constrains).

Background materials for the event will build on our environmental scan and on initial participant feedback from pre-forum outreach. These will frame concrete questions to provide a starting point for discussion, providing common ground and a sense of what has already been done, and expressing the gaps and challenges still remaining. The preparatory materials will include: a

bibliography of published research and other written materials; an environmental scan of existing projects, experiments, implementations, and works in progress; and a set of three brief “design provocations” (one for each of the three main topics) that include a short list of existing work/examples, a paragraph describing the remaining challenges as we see them arise out of prior work, and a set of framing questions.

While this grant will support some participant travel to the Opening Forum, we recognize that not all potential participants have the ability to travel to us. Therefore we will livestream or record the events for public access, and create a Twitter hashtag and online collaborative notes documents to allow for virtual participation. The GRA will be charged with gathering and synthesizing the in-person and virtual feedback we receive during and after the Opening Forum.

The smaller **Two-Day Core Design Group Working Meeting** involves the Core Design Group (CDG, approximately 10 members), and DSG staff. Its goals are to create a work plan for the toolkit development and to develop an initial framework and rough draft of the teaching toolkit. Before this meeting, the co-PIs and Graduate Research Assistant will circulate notes synthesizing the discussion in the Opening Forum Meeting, and the CDG will also include their own responses to the discussion in the initial forum meeting. To provide an initial starting point for discussion, the co-PIs and Graduate Research Assistant will also provide a very general toolkit framework. The teaching toolkit will start with the three main topic areas covered in the Opening Forum.

The final **One-Day Concluding Review Forum** is a large public event with two goals: to expose the draft teaching toolkit and white paper for critique and feedback to a larger audience, and to generate excitement about the project and potential partners to test and refine the toolkit in the future. We again define the audience broadly: practitioners, educators, researchers, and students in LAM and related fields. A draft schedule is included in the supplementary materials.

Evaluation

Evaluation of this project will focus on gathering formative, qualitative input. Evaluation will chiefly serve an internal audience to ensure that we’re meeting the goals of the project and needs of the intended audience; therefore we do not plan on circulating the evaluation results publicly, nor do we plan strongly quantitative assessment.

We will use three methods of formal evaluation. First, we will develop a survey of participants for use after the Opening Forum, with a focus on evaluating (a) how successful we were in ensuring diverse participation, and how we might improve for later events and (b) the logistics and design of the event, particularly whether there was a good balance of time between formal speakers and attendee participation. We will use these results to determine whether we are meeting our diversity goals and how to refine the methods for audience feedback. Second, we will develop a survey of participants for use after the concluding Review Forum, with a focus on evaluating (a) the logistics and design of the event, particularly whether there was enough time or whether there were appropriate methods for broad input into the toolkit and (b) whether there are additional groups to whom we should send the toolkit for review. Finally, we will create an in-depth evaluation instrument for the toolkit, to coincide with the Review Forum, with a focus on evaluation of (a) the content and organization of the toolkit itself, (b) what groups would be most likely to use the toolkit, and (c) channels used for disseminating the toolkit to those groups. We will analyze these survey results, incorporate the resulting information into the final draft, and use their suggestions for outreach and dissemination.

Diversity Plan

Genuinely diverse participation is essential — indeed, definitional — to the success of this project, and we recognize that we need to develop not only a diversity of methods for

participation, but also a diversity of participants. Our goal is to incorporate diverse voices into development and review of the Teaching Toolkit, in support of information and system design for an inclusive National Digital Platform.

We will follow best practices in fostering a diverse pool of participants, using a combination of methods¹⁸. A combination of strong outreach and recruitment targeted at a range of affinity groups, with some participants expressly encouraged to participate, coupled with an open application process (where applications are reviewed blind) will help diversify participation. We will also develop prominent public values statements on inclusion and a code of conduct and make diversity a public goal, using inclusive language on communications to make it clear that we are looking for new voices, and that experience is more important than formal qualifications (e.g., we are looking for participants “with knowledge and expertise to share” vs. only those that are “experts in the field.”) Both public events will be held at no cost to participants, and we have budgeted significant funding for travel support, both for participants at the first Opening Forum and for the Core Design Group members at the second meeting. We will also develop robust processes for public comment and other opportunities for virtual participation for those that cannot travel, including live streaming with synchronous commenting and questioning mechanisms (such as via Twitter or Slack). We will also tap into our Advisory Board’s networks and experience with partnership with diverse groups to help us in this process. To thoughtfully consider the makeup of our participants, we have developed an **Outreach and Recruitment: Initial Scan**, available in the supporting documents.

We seek diversity across many variables, including but not limited to gender diversity; racial, ethnic and cultural diversity; economic diversity; and a diversity of types of institutions and organizations, formal and informal.

Project Resources (personnel, time, budget)

Timeline: December 2016 - December 2018

Budget: Our request for \$99,822 includes subsidies for participant travel, to allow a wide range of attendees, as well as honoraria for 3 speakers at the Opening Forum as well as approximately 10 Core Design Group (CDG) participants. We anticipate that the CDG will provide much of the labor in developing the Teaching Toolkit, and we want to recognize that labor. Our budget also funds a Graduate Research Assistant (GRA) to coordinate events and methods for virtual participation, publicize the project and opportunities for participation, take notes and synthesize feedback, contribute to the development of the Teaching Toolkit, and help prepare the conference proposal submissions that will fulfill the Communications Plan. The rest of the budget is devoted to the logistics of event planning (room rental, food, recording) and a small amount for conference travel to support the Communications Plan; the Northeastern University Libraries will also provide conference travel support.

Northeastern University Libraries Personnel

Co-Principal Investigator (5% time, 24 months): Julia Flanders is the Director of the Digital Scholarship Group in the Northeastern University Library and interim co-director of the NULab for Texts, Maps and Networks; she also serves as the co-director of TAPAS and the director of the Women Writers Project. She has served in positions of leadership in the digital humanities and TEI communities for nearly two decades, serving on the executives of the Association for Computers and the Humanities (president, 2008-2011), the TEI (chair, 2004-2005), centerNet, and the Alliance of Digital Humanities Organizations (Secretary, 2012-present). She has extensive experience developing and managing digital humanities projects and working with

¹⁸ For example, see Milstein 2011 and 2014 as well as “Resources” at the *Who Did You Miss?* project.

sustainable models for digital humanities services, and deep research expertise in digital scholarly communication.

Co-Principal Investigator (5% time, 24 months): Amanda Rust is the Assistant Director of the Digital Scholarship Group. She brings experience in library services for the digital humanities and archives as well as the arts and humanities broadly speaking. She holds office in a national forum for literature librarians, and has presented or written on topics like information literacy and instruction, Drupal and website design, theatre archives, library special collections and Wikipedia, and the digital humanities.

Repository Manager (5% time, 24 months): Sarah Sweeney is the Northeastern University Library's Digital Repository Manager and a member of the Digital Scholarship Group. Sarah brings extensive experience with metadata standards, repository design and implementation, and platforms for digital scholarship. She has led the management of multiple DSG digital projects and is an active participant in the Digital Library Federation and Open Repositories communities.

1 Graduate Research Assistant (to be hired, 10 hrs/week, 15 months) will coordinate events and methods for virtual participation, publicize the project and opportunities for participation, take notes and synthesize feedback, contribute to the development of the Teaching Toolkit, and help prepare the conference proposal submissions that will fulfill the Communications Plan.

Advisory Board

Ellen Cushman (Professor of English and Director of Civic Sustainability, Diversity, and Inclusion), in her work on digital infrastructure for Cherokee language translation, learning, and preservation, has deep experience considering how interfaces for discovery and engagement with Native American documents affect the ways that tribal users and contributors (and also outsiders to that community) encounter those materials.

Moya Bailey (newly hired as an Assistant Professor) has published and spoken widely on the ethical and design challenges of modeling research data drawn from social network data in vulnerable communities.

Giordana Mecagni (University Archivist and head of Northeastern's Archives and Special Collections) works on collecting and sharing archival materials on Boston social justice and community history to a non-academic audience. She has deep ties to and experience developing partnerships with the politicians and community members involved in social justice work in Boston.

Marty Blatt (Professor of Practice in the History department and Director of the Public History program) has a strong background of research and practice in public humanities and issues of community engagement around culturally sensitive topics and materials, with experience in curating and developing exhibits on topics such as "States of Incarceration" (Spring 2018) or "Long Road to Justice – African Americans and the Courts in Massachusetts" (currently in the Edward Brooke Courthouse in downtown Boston).

Élika Ortega Guzman (also newly hired as an Assistant Professor) brings a research focus on global diversity in digital humanities. She writes about digital literature, (not necessarily digital) media, intermediality, materiality, reading practices and interfaces, books, networks, digital humanities, and multilingualism in academia, and is an active member of Red de Humanidades Digitales (RedHD), a network of Spanish-language DH practitioners.

While we have identified an initial Advisory Board, we are open to adding new members once the project begins. We have done work identifying an additional pool of potential avenues for outreach (see our **Outreach and Recruitment: Initial Scan** document), and expect to recruit Advisory Board and Core Design Group members both before and through the Opening Forum. We plan to recruit some CDG members in place before the Opening Forum, but, similar to our

process with the Advisory Board, plan to leave positions open and use the Opening Forum and surrounding publicity as a tool for broad recruitment.

Timeline

A fuller version of this timeline is available in the Schedule of Completion.

December 2016 - March 2017: Confirm participation of existing advisory board; recruit potential participants (attendees, speakers, potential CDG members); issue periodic publicity announcing the project and reporting on progress at intervals; begin hiring process for GRA.

April 2017 - June 2017: Hire and train the GRA; begin logistics for all three events and publicity for Opening Forum; begin preparation for Opening Forum (participant packet, evaluation survey, mechanisms for public feedback).

July 2017 - September 2017: Finish logistics and preparation for Opening Forum; finish recruitment (attendees, potential CDG members) for Opening Forum.

October 2017 - December 2017: Host Opening Forum, including robust note-taking, post-forum evaluation survey, and issue publicity for project website and progress so far; finalize CDG membership, schedule get-to-know-you conference call (11/2017) and in-person meeting (01/2018); prepare CDG Toolkit guidance and other material for CDG in-person meeting; collect and synthesize Opening Forum Feedback, publicize notes, and share evaluation survey results internally.

January 2018 - March 2018: Host CDG Working Group Meeting, including robust note-taking; finalize CDG group structure and work plan; develop draft toolkit materials and draft of white paper with virtual feedback from AB; finalize logistics for Review Forum.

April 2018 - June 2018: Begin content preparation for Review Forum (evaluation survey, live feedback mechanisms, online review method for Toolkit); prepare draft Toolkit and white paper for release at Review Forum.

July 2018 - September 2018: Host Review Forum, including robust note-taking and post-forum evaluation survey; issue publicity for project website and progress so far; release draft Toolkit and white paper for public comment; GRA employment ends.

October 2018 - November 2018: Incorporate public feedback into final Toolkit and white paper; publicize project website, Toolkit, white paper; deliver final materials and reports to IMLS.

Communications Plan

The GRA, together with the co-PIs and Repository Manager, will have primary responsibility for communicating about the project throughout the project lifecycle. The GRA will help library staff put together a robust system for gathering and tracking public feedback -- which will also serve as a measure of audience engagement -- including the creation of a project hashtag and capturing Twitter feedback via Storify. In addition, the GRA will also lead a process of synthesizing public feedback into forms useful for the Advisory Board and Core Design Group. We will livestream forum events, and record them for later review. Finally, we will ask that the Core Design Group charges a member with the task of external communication and outreach, to work in partnership with the co-PIs, Repository Manager, and GRA, and release public progress reports from the Core Design Group where possible.

We have already begun identifying projects, professional communities, and channels to target for outreach, including conferences where we might present on this project. (See the **Outreach and Recruitment: Initial Scan** list, in supporting documents, for details). We will draw on the existing networks of our AB and participating library staff to further enhance our Outreach and

Recruitment list, and actively promote the project via our personal and organizational social media accounts. The Northeastern Libraries provides travel support that will cover one conference per year for professional staff, and is willing to provide additional funding for grant-related presentations at additional conferences.

Finally, we will track project progress on a public blog and share materials developed through the grant, including but not limited to the Open Forum Preparatory Packet, the final white paper, and the Teaching Toolkit. Materials will be preserved and receive permanent URLs in the library's Digital Repository Service (a Fedora/Hydra repository), and all materials will be released for open sharing, re-mix, and re-use, following our general philosophy supporting Open Access publishing.

Part I

A.1 What will be the intellectual property status of the content, software, or datasets you intend to create? Who will hold the copyright? Will you assign a Creative Commons license (<http://us.creativecommons.org>) to the content? If so, which license will it be? If it is software, what open source license will you use (e.g., BSD, GNU, MIT)? Explain and justify your licensing selections.

All materials will be owned by Northeastern University, and will be published under a Creative Commons Attribution license (CC-BY) to permit the maximum circulation and reuse. The purpose of all materials created is to prompt further discussion and development of derived materials, so we will not impose any restrictions on derivatives; in the digital humanities community, significant arguments have also been made that prohibiting commercial reuse is undesirable and unnecessary (see for instance <http://nowviskie.org/2011/why-oh-why-cc-by/>).

A.2 What ownership rights will your organization assert over the new digital content, software, or datasets and what conditions will you impose on access and use? Explain any terms of access and conditions of use, why they are justifiable, and how you will notify potential users about relevant terms or conditions.

All materials will be published without restrictions on access or use.

A.3 Will you create any content or products which may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities? If so, please describe the issues and how you plan to address them.

We only plan on videotaping presentations only, not the small-group discussions or working meetings. We will gather appropriate releases from the public presenters. We will notify the audience that videotaping is occurring, and position the cameras so that audience members are generally not captured. We will also create mechanisms for those that do not wish to be videotaped or photographed to indicate those wishes so that we can abide by those wishes. We will let online Twitter discussants know that their contributions will be captured via Storify, and where we wish to widely broadcast comments we will contact the individual in question and ask if they prefer to remain anonymous or credited.

Part II

A.1 Describe the digital content you will create and/or collect, the quantities of each type, and format you will use.

The content created will include three different kinds of materials:

1. Video footage of public forums: roughly 20-24 hours of footage. Footage will be stored using H.264 video and AAC audio using an .mp4 wrapper.
2. Public notes, draft documents, and the final white paper: roughly 10-20 documents. These documents will be created as Google documents (to support collaborative editing

and public commenting) and then saved as both rich text and HTML (for long-term storage) and PDF (for final dissemination).

3. The curricular materials that make up the teaching and learning toolkit: roughly 20-30 documents. These documents will be created as Google documents (to support collaborative editing and public commenting) and then saved as both rich text and HTML (for long-term storage) and PDF (for final circulation).

A.2 List the equipment, software, and supplies that you will use to create the content or the name of the service provider who will perform the work.

The video footage will be captured with whatever videotaping equipment is customarily used by Northeastern University's videotaping service. Because the purpose of the videotaping is to provide a record of the forum events for those who cannot be physically present (and for subsequent reference), the quality of the video data is only important insofar as it serves that function; it is not high-value data in itself. There are several different options at Northeastern for performing this service and we will determine which to use closer to the time of the event; all of the options we are considering will use industry-standard equipment that will capture data of sufficient quality for purposes of this project.

No specialized equipment, software, or supplies are needed for the textual materials (drafts, notes, curricular materials, white paper). These documents are probably not "digital content" in the sense ordinarily intended by this form. All of the textual content described here will be created by the project participants during the course of the project using their own computers and collaborative authoring tools such as Google Docs.

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to create, along with the relevant information on the appropriate quality standards (e.g., resolution, sampling rate, or pixel dimensions).

The digital file formats to be created are: RTF, HTML and HTML5, MP4, and PDF. For the textual materials being created (for which the file formats are RTF, HTML, and PDF), quality standards are not relevant. For the video materials, the video will be captured ideally at 1080p, though 720p will be accepted, depending on equipment], which will be scaled down by the streaming server to accommodate a wide range of connection speeds at the user end. Audio will be captured and streamed at at least 256 kbps.

B.1 Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

The textual materials created under this grant (public drafts, white paper, curricular materials) will be carefully proofread by the graduate research assistant. The quality of the video will be monitored as part of the videotaping service, and the edited video footage will be reviewed by the project team for clarity and appropriateness. Passages of discussion may be redacted at the request of the participants involved.

B.2 Describe your plan for preserving and maintaining digital assets during and after the award period of performance (e.g., storage systems, shared repositories, technical documentation, migration planning, commitment of organizational funding for these purposes). Please note: You may charge the Federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the Federal award. (See 2 CFR 200.461).

All of the digital content created under this grant will be deposited into Northeastern University's Digital Repository Service (DRS, <https://repository.library.northeastern.edu>), and will be disseminated from there. The DRS is a Fedora/Hydra repository that was built to maintain and preserved scholarly, archival, and administrative digital objects produced or acquired by Northeastern University. Northeastern University Libraries has enthusiastically committed to supporting the DRS as a storage and preservation tool and as an essential library service. The technological components that comprise the DRS are spread over three servers, all supplied and maintained by the university's Information Technology Services (ITS). ITS provides remote storage in two different geographic locations, as well as frequent backups of DRS content. DRS staff perform regular audits of DRS object metadata, which include validity and integrity checks. The formats used for these materials all fall under the DRS's standard ingest, migration, and streaming protocols. Because of the small quantity of data and standard formats involved, additional resources required for the long-term preservation of these materials in accessible form are effectively nil.

C.1 Describe how you will produce metadata (e.g., technical, descriptive, administrative, or preservation). Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, or PREMIS) and metadata content (e.g., thesauri).

Administrative, technical, and descriptive metadata for each item deposited in the DRS will be produced at the time of upload. Technical metadata will be generated using information gathered from the file itself and stored using FITS. Descriptive metadata will be created using the standard metadata creation form for repository ingestion and stored using MODS. While MODS is the primary metadata standard for DRS objects, Dublin Core is also used to satisfy minimum Fedora requirements. Because these materials do not form a large-scale collection and their successful dissemination depends on the circulation of links rather than on discoverability via a search interface, detailed metadata does not play a key role in the usage of these materials.

C.2 Explain your strategy for preserving and maintaining metadata created and/or collected during and after the award period of performance.

Metadata is not a central component of this project; it serves as a small descriptive adjunct to the specific content items (documents, video) being created. The metadata for each item will be preserved in the DRS along with the content items themselves, and will be curated and migrated as part of the DRS's regular migration and maintenance activities.

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of digital content created during your project (e.g., an API (Application Programming Interface), contributions to the Digital Public Library of America (DPLA) or other digital platform, or other support to allow batch queries and retrieval of metadata).

As noted above, the materials created for this project do not constitute a large-scale data collection, nor are they in genres that lend themselves to aggregation within other standard collections, since they are not digitized versions of primary-source content. Metadata for DRS objects flagged as public will be made available through the DRS's API, OAI-PMH harvesting, and through indexing by Google and others indexers. However, our most effective avenues of dissemination will be outreach to specific audiences via publications, discussion venues, social media, and conferences.

D.1 Describe how you will make the digital content available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

All materials created under this grant will be published on an open-access basis via the DRS, and links to these materials will be publicized via links from the Digital Scholarship Group web site, which in turn will be disseminated via social media, discussion list postings, and conferences/publications. The DSG will create a website for this initiative where we will provide a descriptive narrative about the forums, including information about participants, drafts for comment, suggested readings, and other contextual materials. From this site, we will provide links to the video footage and final versions of all documents

D.2 Provide the name and URL(s) (Uniform Resource Locator) for any examples of previous digital collections or content your organization has created.

<http://arader.library.northeastern.edu>,

<https://repository.library.northeastern.edu/collections/neu:rx913q62j>

<http://aproudpast.library.northeastern.edu>,

<https://repository.library.northeastern.edu/sets/neu:cj82m8833>

Original Preliminary Proposal

National Forum on Culturally Sensitive Practices and Technology

Proposal Description

The acquisition, housing, and care of cultural objects has always had colonial implications. The library, archives, and museum (LAM) world has begun to grapple with the consequences of asserting control over items removed, often forcibly, from their original cultural context. Professional LAM communities now assert the desire for genuine, responsive partnership with the communities in which these artifacts were created, deeper understanding of the hegemonic role of knowledge representation via standardized ontological decisions, and investigation into the role of algorithm, interface, and tool design in reinforcing power differentials inherent in the status quo.

These questions are inherent in the turn to the digital and exacerbated by the renewed focus on community archives, both within the LAM world and without. Increasingly, researchers across all disciplines seek to capture and preserve the voices of communities not often given a voice, communities which have often experienced forms of trauma. Now that academic researchers create and theorize “the archive” at an ever-increasing pace, what are the responsibilities of the researcher, the designer, the curator, or the archivist in creating new open digital collections ready for automated harvesting, software-assisted analysis, and broad re-use and reinterpretation? What is the relationship between knowledge representation, naming, and power in relation to not only interactions with marginalized communities but also the design of tools and interfaces? When tools and interfaces limit or constrain interactions with documents and items, are they responsive to community needs or do they truncate ways of knowing? What commonalities exist across communities of users, and what do not?

We seek funding for a National Forum to begin to answer those questions. The forum events will bring together not only librarians, archivists, and curators, but also historians and oral historians, humanists and digital humanists, programmers and metadata experts, anthropologists and sociologists -- all people with a vested interest in the ethical creation and support of digital archives involving marginalized communities. Most importantly, these events will include members of marginalized communities who are seeking themselves to create their own digital archives, perhaps with a larger institutional partner, perhaps not.

These events will draw on both local and national expertise. Locally, we have partners with experience in working with communities that have experienced trauma (oral historians in *Our Marathon: the Boston Bombing Digital Archive*), tribal communities to preserve indigenous languages (Ellen Cushman), and ethical research partnerships with marginalized communities (Moya Bailey). We expect to invite additional participants involved in: culturally sensitive design (e.g. the Mukurtu CMS); developing copyright and re-use licenses sensitive to the needs of indigenous materials; ontology and data modeling, including cultural heritage groups within the standards community (e.g. CIDOC CRM); archives, libraries, and data and other curation at a range of large and small cultural institutions; and community work (outside the traditional LAM environment) collecting and digitizing at-risk heritage materials. Potential keynote speakers and discussion leaders include Wendy Chun, Élika Ortega Guzman, Hope Olson, Angel Nieves.

Relevance to Agency Priorities

This proposal addresses themes that recur across both IMLS' National Digital Platform and Learning in Libraries agency priorities: inclusive, radical collaborations to increase diversity and community engagement. All three reports from the 2015 IMLS Focus Conference (*Engaging Communities, Learning in Libraries, National Digital Platform*) paint a vivid picture of partnership with local communities in support of participatory learning, digital literacy and broad inclusion. We seek to develop best practices and professional development to help librarians, archivists, curators, and researchers

incorporate ethical community partnership in their daily work, and recognize the often hidden organizational and emotional labor needed to maintain mutually beneficial partnerships.

Potential Impact

Community engagement, recovering history, and collecting and archiving the voices of marginalized groups are all topics of the moment, both within and without the LAM community, where important and relevant work has already been supported by the IMLS (e.g. the recently funded *Digital Atlas of Native American Intellectual Traditions*). While previous work in this area exists, best practices have seldom reached cross-disciplinary prominence. This National Forum series will produce a white paper and set of best practices developed with very broad and diverse input, including a robust scan of prior work, that will serve as a resource for further interdisciplinary communication. We will also produce a preliminary curriculum for librarians, archivists, curators, and researchers working with sensitive digital material, for use in library and information science education or as a foundation for a future grant application to the IMLS Laura Bush 21st Century Librarian program supporting professional development within the field.

Projected Goals and Outcomes

We will plan and host two meetings over one year. Each meeting will be a two-day event with small groups of 20-25, with the second half of the second day (keynote, panel, and open discussion) open to the public. Participants will be chosen through an application process to ensure diverse representation and a wide range of expertise. These meetings will consider the issues around naming, knowledge representation, and power in two areas. *Ethical Partnerships with Marginalized Communities* will focus on practices and considerations in building strong, ethical relationships surrounding the collection and curation of culturally sensitive archival materials and data. *Cultural Norms in Interface and Tool Design* will examine design considerations for data curation tools and interfaces that support more diverse ways of knowing and diverse encounters with archival materials. The outcomes of this National Forum series will be:

- White paper with environmental scan and best practices
- Preliminary curriculum for future LAM professionals
- Identification of partners to use and/or further develop curriculum
- Identification of organizations willing to convene groups to develop and host relevant standards
- Submission of conference presentations based on white paper to appropriate LAM and other disciplinary conferences (American Library Association, Society of American Archivists, Digital Humanities, American Studies Association, etc.)

Budget

Our request for \$99,680 includes subsidies for participant travel, to allow a wide range of attendees, as well as a graduate research assistant to coordinate events, take notes, contribute to curriculum development, and prepare conference proposal submission and final write-up of white paper and curriculum.

- Participant travel subsidies: \$48,000 (\$24,000 per event: 20 participants receiving \$1,200 each)
- Speaker honoraria: \$6,000 (\$3,000 per event: 3 speakers receiving \$1,000 each)
- Speaker travel and lodging: \$9,000 (\$1,500 per speaker, 6 speakers total)
- Event costs: \$6,000 (\$3,000 per event: refreshments, room rental, outreach materials)
- Graduate Research Assistant: \$28,500 (full-time for academic year and summer) and benefits at 7.65% (\$2,180)
- Indirect costs waived by Northeastern
- Final total: \$99,680