

**Abstract**

The National Forum on Ethics in Web Archiving will be a three-day national forum that convenes archives, museums, and libraries professionals, community representatives (socially engaged artists, activists, and members of community organizations), and interdisciplinary practitioners (social media researchers, lawyers) to discuss the ethical challenges, both technical and social, involved in archiving the web and social media. Following the forum, a best practices white paper and proceedings will be published.

The project is led by Rhizome in collaboration with the University of California at Riverside Library (UCR), the Maryland Institute for Technology in the Humanities (MITH), and the Documenting the Now project.

Rhizome is a non-profit organization that supports contemporary art that creates richer and more critical digital cultures. An affiliate in residence at the New Museum in New York, Rhizome is the leading international born-digital art organization, online since 1996. Since 1999, Rhizome has developed collections of digital art, and pioneered digital preservation strategies for the maintenance of these forms of. Its Webrecorder platform, developed with the support of the Mellon Foundation, aims to give users powerful tools to create their own decentralized web archives.

Documenting the Now is a tool and a community developed around supporting the ethical collection and use of social media content. A project of University of Maryland, University of California at Riverside, and Washington University in St. Louis, it is supported by the Mellon Foundation.

The forum is intended to address the need for ethical standards for web archiving in the age of social media that can be adopted by professionals as well as community-based users. Archivists need better frameworks for ethical content collection, so that they can safeguard the communities whose histories they aim to document. Likewise, users need to be kept safe; the content they publish on the web and social media collections should not be used by public archives in ways that cause them harm.

The program will address this need by bringing together communities, archivists, scholars, developers, and designers to establish shared goals and methods that will lay the foundation for richer, more representative, non-oppressive web archives. These goals will be distributed after the event in the form of a proceedings publication for specialist audiences and a short white paper, written for users at all levels of expertise. Through these activities, the program aims to lay the foundation for web archives that will better serve the American public and the historical record.

## **Narrative**

This National Forum is intended to address the need for ethical standards for web archiving in the age of social media that can be adopted by professionals as well as community-based users.

The program will address this need by bringing together **community representatives** (activists, socially engaged artists, and representatives of community organizations who work to improve the lives of people from a particular neighborhood or social group, and are empowered to speak on their behalf), **archives professionals** (those working in the museums, archives, libraries field), and **interdisciplinary practitioners** (scholars, social media experts, designers, web developers) for a structured conversation to pool knowledge and establish shared goals and methods that will lay the foundation for richer, more representative, non-oppressive web archives—archives that will better serve the American public and historical record.

### **1. STATEMENT OF NEED**

*What nationally significant challenge does your proposal address? Why is the issue your proposal focuses on nationally significant and timely?*

The dramatic rise in the public's use of web and social media platforms to document events of historical significance presents archivists with a tremendous opportunity to transform appraisal, collection, preservation, and discovery of this new type of research resource, and to involve the public in this process in new ways.

Over the past twenty years, web archiving technology has improved extensively. From the early 1990s with the pioneering Internet Archive and their Wayback Machine, to newer tools such as Rhizome's Webrecorder system, web archiving and the technical expertise around developing and improving tools have allowed for better high-fidelity archives of the personalized, dynamic web.

While the promise of web data collections is enormous for archives, the scholarly community, and the public at large, the privacy and safety risks to public users of these platforms are significant. Individual organizations have made relatively isolated efforts to address these issues, but they have not been given adequate attention in the libraries, archives, and museums field. Ethical issues in web archiving are exacerbated by gaps in web and social media platform literacy; rapid changes in technology; widespread use of the web for public shaming and other kinds of abuse; and increased opportunities for surveillance, especially for individuals taking part in public displays of civil disobedience.

*How does this proposed project differ from, complement, or build upon previous work in this area?*

No ethical guidelines for web archiving yet exist, and conversations about ethics in web archiving are at a nascent stage. To address this, the National Forum on Ethics and Web Archiving will convene archivists and users from a range of backgrounds to pool knowledge and experience on these topics. This forum will provide an opportunity to lay a foundation for more ethical web archives that also reflect the urgency and diversity of digital culture today. The Forum will foster a deep understanding of these ethical landmines and a commitment to non-oppressive content collection, while emphasizing the richness that social media can bring to our understanding and recording of contemporary culture.

Over the past year, the Mellon Foundation-funded Documenting the Now project (DocNow) has been addressing some of the issues surrounding best practices in collecting and preserving social media for public archives, focusing in particular on social media activism. Issues that have arisen include requests by law enforcement agencies for access to Twitter archives, and re-publication of social media posts by news media without a user's consent. In response to these issues, DocNow has run design workshops in order to identify ethical issues in social media archiving and possible design approaches to address these inequities. This process surfaced important issues that demand further investigation, highlighting the need for a wider conversation

about ethics in web archiving. In particular, it highlighted the need to look at design approaches in parallel with questions of methodology and outreach.

Rhizome has been exploring similar questions through Webrecorder and its other digital preservation initiatives. The Webrecorder system allows relatively untrained users to easily create high-fidelity archives of social media and embedded content, thus enabling decentralized and vernacular web archives. By putting powerful web archiving tools into the hands of users, Webrecorder creates the possibility for richer, more inclusive web archives to be built, but it also creates new possibilities for harm. Rhizome's attempts to understand and grapple with these issues have also highlighted the pressing need for a broader, more structured conversation that looks beyond any one tool or initiative to consider the web archiving field as a whole.

The forum will also build on work done in the field of sociology to develop ethical frameworks for decision-making about the use of social media content in scholarly research. One particularly useful precedent is the publication "Social Media Research: A Guide to Ethics" (Townsend and Wallace, 2016). The National Forum on Ethics in Web Archiving will be an ideal opportunity to bring this research into dialogue, connecting it with the archives field as well as a wider, interdisciplinary group of professionals and community members to discuss and develop technical solutions and some guidelines for ethical decision-making regarding the collection of social media and web data.

*Specifically, how will your project address the issues identified in the project category you have selected for this program (as described above in Section A)?*

This project aims to meet the IMLS agency-level goal of "strengthening museums and libraries as essential partners in addressing the needs of their communities."

The communities served by museums and libraries around the world increasingly use online platforms as a crucial part of their social, cultural, and political lives. Thus, memory institutions face serious dilemmas in collecting and safeguarding knowledge on behalf of communities. If museums and libraries choose not to archive social media, they are seriously limiting the kind of knowledge they are able to gather. If they do, they run the risk of bringing unwanted attention or even harm to the communities they aim to serve.

The National Forum on Ethics and Archiving the Web aims to bring together community representatives and archives professionals to pool knowledge and draft ethical guidelines in order to strengthen museums' and libraries' ability to create archives of personal web content in ways that allow for richer and more inclusive representation of their communities without putting them in harm's way.

This project aligns closely with several IMLS National Digital Platform key goals as they are described in the 2015 IMLS Impact Report on the National Digital Platform.

The IMLS key goal of "Championing Diversity and Inclusion" aims to ensure that the national digital platform serves all Americans. This aligns with the aim of this initiative to ensure that more Americans have more of a say in how, or indeed whether, they are represented in national web archives. It aims to do this by bringing small institutions and community-based archivists, who have the local knowledge required to build more representative and ethical web archives, together with representatives of larger institutional archives, who have more access to preservation infrastructures to share knowledge. A set of guidelines developed as a result of this conversation will go on to inform archival practice at a range of scales.

This also supports the IMLS key goal of "Engaging Communities," which argues that users should not only be thought of as beneficiaries of the national digital platform, but as its contributors. The EAW program will include a special focus on how web archiving can be conducted in a decentralized fashion without exposing

users to greater risks of privacy or other incursions. It will also directly involve community users in the program as presenters and invited respondents.

Finally, the program supports the IMLS key goal of “Enabling Technology,” which identifies interoperability as a key technological concern facing communities. Interoperability is particularly crucial to facilitating a culture of community- and user-led web archiving, because it allows grassroots efforts to be connected together in a larger national archive. The EAW program will support this effort by involving developers and designers who are actively involved in developing web archiving tools such as Perma.cc, Archive It, DocNow, and Webrecorder. Conference sessions will draw on and inform their work, and a day of unconference workshops will allow them to participate in testing out new ideas for decentralized archiving.

## 2. **PROJECT DESIGN**

This project will take place from October 1, 2017 to September 30, 2018. It will be co-organized by Rhizome and DocNow, under the leadership of Project Director Michael Connor.

### *Goals*

The overarching goal of this project is to facilitate the creation of richer, more inclusive web archives that protect the safety and privacy of users.

This includes the following sub-goals:

- To bring together an interdisciplinary group of archives professionals, community users, researchers, and web archiving tool developers and designers for a conversation about shared goals, pitfalls, and best practices;
- To inform the development of web archiving tools with community needs and ethical considerations in mind;
- To create connections between grassroots community efforts and larger institutional efforts in web archiving;
- To foster collaboration among decentralized web archiving efforts;
- To inspire the creation of web archives that are responsive to community perspectives and needs.

### *Outcomes*

The central outcomes of the project will be a conference, published proceedings, and a two-page white paper that will inform web archiving practice.

The program will:

- Convene practitioners from a wide range of backgrounds (professionals, community representatives, and other interdisciplinary practitioners such as designers and scholars) for an event at the New Museum in New York City;
- Offer individuals and groups engaged in community-centered web archiving a platform to share knowledge, experiences, hopes, and fears;
- Facilitate dialogue between community and institutional, formal and informal web archives;
- Draw on applicable knowledge from other fields, especially social sciences;
- Connect developers and designers of web archiving tools with a broader range of users;
- Establish basic principles and approaches to ethical web archiving, and raise awareness of these among practitioners engaged with web archiving.

### *Assumptions*

A central assumption in this program is that web archiving has a responsibility to represent American internet culture in its richness and diversity. In particular, there is a pressing need to ensure that the stories of activist

groups and community organizations are represented in the historical record—and social media is often a big part of this story.

However, web archiving can and does cause harm to the communities it aims to serve. This assumption has been borne out by the experience of Documenting the Now in maintaining archives of tweets related to the Black Lives Matter movement, particularly with respect to the use of social media as evidence of criminality. There are also numerous recent examples in which archive platforms have been used to host web content used for the public shaming of other users. There is a need for a better understanding of the potential harm caused by web archives and the steps that can be taken to lessen or avoid such harm.

The potential harm caused by web archiving can be mitigated by raising public awareness, building networks of collaborative practice, and identifying design approaches to inform the development of tools that facilitate ethical collecting.

A further assumption is that an ethical web archiving culture will require greater collaboration among tool-builders, communities, and institutions. While not every ethical problem in web archiving can be solved through design, those responsible for creating web archiving tools have a responsibility to consider the safety of users in their design process. Communities are essential to ethical web archiving, as they have the local knowledge needed to avoid causing harm, while ensuring that community perspectives are represented. And institutions are crucial to ethical web archiving as providers of resources, knowledge, and infrastructure.

#### *Risks to the project*

Communications pose a central risk to this project. As the main outcome is conversation and awareness-raising, if a communications plan is not developed and executed properly, the event, the livestream, and the published white paper and proceedings will not reach its audience, and the outcome of the program will not be achieved.

This risk will be mitigated by developing a communications plan, which will be managed by Rhizome with the advice and support of the New Museum. This plan will leverage the community and institutional contacts developed by Documenting the Now and Rhizome's Webrecorder project. Further, a network of advisors will be convened in the conference planning process, and their networks will be drawn on to further propagate information about the program. Progress against the communications plan will be checked weekly as part of event planning meetings.

Project organization and management is another potential risk. Event planning and logistics will be a major undertaking, involving group decision-making and liaison with numerous partners and participants. Miscommunication or poor planning will lead to suboptimal results and failure to achieve the project's objectives.

To mitigate this risk, Rhizome's artistic program team, led by Artistic Director Michael Connor, will lead the event planning effort, bringing extensive past experience in managing complex events including the annual *Seven on Seven* conference. As the event takes shape, Connor will work closely with the New Museum staff—including public programs, special events, technical, ops, security, and front of house departments—to ensure that the event runs smoothly.

Finally, there is a risk that the white paper and proceedings will not be finished. Here, the central risks are contributors not filing their work, and a failure to come to collective decision about the contents of the white paper.

Rhizome has extensive experience in running a deadline-oriented publishing program, with programs like its weekly series Net Art Anthology. Signing contracts with contributors, setting deadlines, and following up in a timely fashion tend to mitigate many of the risks involved in publishing to deadline.

*What specific activities will you carry out?*

The National Forum on Ethics and Archiving the Web will be a three-day symposium including panel presentations, keynotes, hands on workshops. Conference sessions will be livestreamed, and these will be archived online for continued access.

Participants could include **archives professionals** such as Steven Fullwood (Associate Curator at the NYPL's Schomburg Center); Ilya Kreymer (founder of Webrecorder, and senior developer for Rhizome), Dr. Safiya Noble (assistant professor at the USC Annenberg School of Communication and co-founder of the Information Ethics & Equity Institute), Jefferson Bailey (Internet Archive), Jarrett Drake (Digital Archivist at Princeton), and Stacie Williams (Team Lead, Digital Learning & Scholarship, Case Western Reserve University); **community representatives** such as Guadalupe Rosales (artist and founder of the *Veteranas & Rucas* Instagram account), Keith Calhoun and Chandra McCormack (founders of the L9 Art Center in New Orleans), Mariame Kaba (@prisonculture on Twitter, organizer, educator, and curator), Lauren Chief Elk (activist and domestic violence victim advocate), and Cathy Cohen (David and Mary Winton Green Professor of Political Science and the College at the University of Chicago, and founder of the Black Youth Project); as well as **interdisciplinary practitioners** such as Zeynep Tufekci (sociologist and associate professor at UNC-Chapel Hill); Dr. Leanne Townsend (sociologist with expertise on social media research ethics and rural creative economies), and Jane Anderson (legal expert on copyright and fair use as they pertain to indigenous cultures).

Following the symposium, a white paper will be produced by the organizers. This will be a simplified executive summary of recommended practices, goals, and methods for ethical, community-focused web archiving. It will be written in non-specialist language and feature approachable visual design. The white paper will be distributed as a PDF.

This will be accompanied by a proceedings publication, also distributed as PDF, in which contributions from participants will offer more first-person knowledge of the risks and opportunities presented by web archiving in the age of social media.

*How are they informed by current theory and practice?*

The conference will draw together leading archives professionals, community representatives, and related interdisciplinary practitioners, and will feature conversations, case studies, workshops, and keynotes that capture the current discourse and practice surrounding web archiving.

*When and in what sequence will your activities occur?*

October 2017: Co-organizers and advisory committee will convene via conference call in to identify potential guests and set panel topics.

November 2017: Invitations will be sent to invited speakers, and open call for funded participation will be announced.

Early December 2017: Submissions reviewed and responses issued.

March/April 2018: A three-day forum will take place in late March 2018 at the New Museum in New York City. It will include traditional panels and keynotes as well as unconference sessions and a design workshop on ethical data archiving tools.

April 2018: Co-organizers will review conference outcomes and circulate an outline of the white paper.

May 2018: Deadline for proceedings contributions from conference participants and advisors. Deadline for feedback on white paper from participants and advisors.

September 2018: Publication and promotion of proceedings and white paper with recommendations for ethical web and social media archiving.

*How does the project design allow for input, consensus building, and buy-in from others in and/or outside the field?*

The planning will take place with an advisory committee who will be asked to identify likely guest speakers and invitees, and to review and comment on the program as it develops. This would include library and archive professionals, community organizers, and activists.

A microsite and open call will be published as well, inviting archivists to apply for funded participation as speakers or conference attendees. This will especially prioritize those working without institutional support. This will follow the model of Rhizome's successful Internet Art Microgrants program, and leverage the extensive professional and social media networks established by Documenting the Now, Rhizome, and the New Museum to solicit applications.

The forum will be livestreamed and Twitter interaction will be highly encouraged during all panels and presentations. While standards and guidelines work will begin at the forum, they will be completed post forum with significant community input including a period of open feedback on all proposed documentation.

*Who is the audience for the project and how will they participate? If applicable, how will you meet the needs of underserved groups or communities?*

The audience for the forum are community representatives with an interest in social media's role in telling the story of their community, museums, libraries, and archives professionals, interdisciplinary practitioners from social sciences and web technology who make use of archiving tools, and web and social media users who are interested in the internet's cultural heritage.

We are specifically interested in participation from members of traditionally marginalized communities in the United States who are interested in how they may be represented in archives of the web, including LGBTQ, African American, Hispanic American, Women, and people living in economically disadvantaged rural and urban communities, among others.

Participation will take place via in-person attendance at the conference. Grant funding will be allocated to funding participation by attendees from nontraditional institutions.

Further participation will take place online via livestream and social media discussion surrounding the event, and by accessing the conference videos, proceedings, and white paper after the fact.

*If applicable, how will you meet the needs of underserved groups or communities?*

The project will meet the needs of underserved groups and communities by developing ways that their histories can be reflected in web archives so that these records do not cause incursions of privacy or bring harm to those archived.

Representatives of underserved communities will be given priority to receive funding for travel and attendance at the conference, giving them the opportunity to gain and share skills and knowledge, and to build relationships within the archives community.

*Who will plan, implement, and manage your project?*

The project director will be Michael Connor, Artistic Director of Rhizome. Connor has extensive experience in organizing events on a national scale, including the *Seven on Seven* conference which is held annually at the

New Museum. The **co-organizing team** will include Connor, Bergis Jules, University & Political Papers Archivist at UC Riverside and Community Lead, DocNow, and Ed Summers, Lead Developer at Maryland Institute for Technology and Technical Lead of DocNow.

A six-person **advisory committee** will be assembled, including community representatives, libraries and archives professionals, and web or social media specialists. The advisors will be asked to provide suggestions and feedback during the planning process.

Further organizational support will be offered by **Rhizome staff**, including Aria Dean, Rhizome's Assistant Curator, Kaela Noel, Rhizome's Editorial Manager, and Dragan Espenschied, Rhizome's Preservation Director.

**New Museum staff** will offer assistance and support in the form of strategic marketing and communications assistance, as well as technical and operational aspects of staging the event itself.

*What financial, personnel, and other resources will you need to carry out the activities?*

The project will require the following resources, which will be met by the project team listed above and by outside contractors as indicated.

#### Event planning and outreach

- Planning resources - desk research, online collaboration, conference calls [co-organizing team]
- Website design and development [Rhizome staff and outside contractors]
- Promotion of call for funded participation [Rhizome staff, co-organizing team, advisory committee]
- Speaker invitations sent [Rhizome staff, co-organizing team]
- Review of applications to attend and speak at the event [co-organizing team, advisory committee]
- Design and production of conference brochure [Rhizome staff and outside contractors]
- Promotion of conference via press, emails, social media, advertisements, etc. [Rhizome staff and New Museum staff]

#### Event staging

- Travel, lodging, hospitality, and stipends for invited speakers and funded attendees
- Guest liaison
- Use of the New Museum venue and event spaces for three days (in-kind, confirmed)
- Conference technical staff and livestream
- Security, maintenance, and front of house staff [New Museum staff]
- Social media management [Rhizome staff and New Museum staff]
- ASL interpretation [Outside contractors]

#### Post-event outreach and publication

- Transcription [Outside contractors]
- Solicitation of feedback [Co-organizing team and Rhizome staff]
- Coordination, editing, design, and production of white paper and proceedings [Rhizome staff and outside contractors]
- Promotion of publication and white paper [Rhizome staff, New Museum staff, co-organizing team, advisory committee]

Rhizome requests \$100,000 for the three-day symposium. An additional \$8,000 of support will come from the Knight Foundation to cover cost for hospitality and event production, and in-kind support of \$18,000 will come from Rhizome and the New Museum towards staffing, marketing and PR, and usage of the venue. The allocation of these funds is described in the Budget Justification.



*How will you track your progress toward achieving intended results? How will you include evaluation and performance measurement in your plan?*

Meetings and project management tools will be used for ongoing project tracking. Progress on all aspects of the project will be tracked on a GitHub repository for the conference. Program development and outreach will be tracked in a biweekly conference calls for co-organizers, and publishing progress will be tracked in Rhizome's biweekly editorial meeting. Progress toward ensuring that conference participants adequately represent a nation of internet users will be tracked via period check-ins with advisors. Press and communications progress will be tracked via Rhizome's biweekly press & communications briefing with the New Museum.

The first opportunity for formal performance measurement will be the response to the open call for funded participation. With forty available slots, the open call should receive upwards of one hundred applications from potential participants of diverse backgrounds.

At the end of each day of the conference, a wrap-up session with conference organizers and advisors will give an opportunity to assess the project in progress, address any issues arising, and pick up on important learning as it happens.

All conference participants will be asked to complete a survey at the end of the event as a way of gathering initial feedback on progress toward the stated project goals. Several weeks later, a draft white paper with ethical guidelines for web archives will be circulated, which asks for feedback and assessment against stated project goals.

*How and with whom will you disseminate your project's results?*

Project results including any meeting report, guidelines, standards, or tools will be made freely available on the project website and also through the Rhizome and Documenting the Now sites. Guidelines and standards will also be shared with the Society of American Archivists for inclusion in the resources section of their website. Project results will be shared with participant network as well.

The project results will be presented in a format where they can easily be downloaded and reused.

*How will you sustain the project beyond the funding period (website, panel videos, tweets about the forum)?*

The circulation of the white paper and proceedings is expected to spark further conversation, with Rhizome and Documenting the Now continuing to facilitate and lead this area of research and discussion as an ongoing program activity.

The website for this event will serve as an ongoing archive, including stable versions of panel videos and social media discussions surrounding it.

### **3. NATIONAL IMPACT**

*What national impact will the proposed project have on the library and archives field?*

The personal nature of documenting participation in historical events via the web and social media present archivists and researchers with new and exciting opportunities to engage with the data generated by individual users. While the promise of web data collections is enormous for archives and the scholarly community, the risks associated with collecting and preserving that type of content for the long term are equally important but have not been given adequate attention.

There have been no field-wide efforts to address the ethics of archiving the web. This lack of action has left major gaps in a.) ethically-based standards development and tool design, b.) identification of vulnerable communities, and c.) addressing how a lack of ethical considerations can specifically impact lives.

This will be the first time a major meeting will take place specifically focused on ethics of archiving the web in the United States, and including archives professionals, community representatives, and interdisciplinary experts.

This program will help raise national awareness of potential ethical issues in web archiving, and will form the basis of a set of ethical standards for building archival collections from the web. It will also identify ways in which vulnerable communities can be harmed for lack of dealing with these issues, and propose principles that can be applied in the design of ethical web archiving tools and collection-building strategies.

This will ensure that the archives and libraries community is taking steps to protect the privacy, content rights, and safety of users whose data is being collected, preserved, and made accessible in public repositories.

*How does the project effectively address current issues that concern the library or archive fields and will have a lasting impact on the field(s)?*

Currently there is a lack of resources, standards, and guidelines to address ethical issues around archiving web data to build research collections. Focusing on these ethical issues through the Forum and white paper, including identifying how traditionally marginalized people can be harmed by unethical web data archiving, will address an area that cultural heritage professionals have identified as a need.

During the work of the Documenting the Now project, including their extensive work engaging with the public and gathering community feedback, archivists and others in the cultural heritage sector consistently referenced the difficulty in planning for the ethical considerations around archiving on the web and have frequently called for resources and guidance that can assist in appraisal and collections development decisions. The National Forum on Ethics in Web Archiving will address archivists concerns around the ethics of archiving on the web by bringing together professionals and community members committed to designing and developing guidelines, standards, and resources to address these issues.

*Referring to your Statement of Need, how will your project's results address the need, problem, or challenge you have identified? These may be in addition to, but not instead of, the Performance Goal(s) and Performance Measure Statement(s) referenced above.*

This project will strengthen museums and libraries as essential partners in addressing the needs of their communities by bringing together professionals from these institutions as part of a larger, interdisciplinary group that aims to surface communities' needs with regard to web archiving – including the sometimes competing needs for privacy, safety, and representation – and address them in a set of best practices to inform the wider field.

*Referring to the Performance Goal and Performance Measure Statement(s) selected on the Program*

*Information Sheet prepared for your application, how will you collect and report the corresponding data?*

The data for assessing this project will be collected via questionnaires distributed at the end of the conference, and a second questionnaire distributed online.

*What are the project targets for these performance measures? How will you measure success toward these targets?*

The National Forum on Ethics in Web Archiving should attract 100 participants and attendees, of which 50% should be from the museums, libraries, and archives profession, and the other 50% community representatives (artists, activists, or members of community organizations) and interdisciplinary practitioners.

A further 3,000 people should watch the livestream, and the white paper and proceedings should reach 5,000 people, as measured by web analytics.

Our aim would be for 50% of conference attendees representing the museums, archives, and libraries field to answer affirmatively to at least one of these questions:

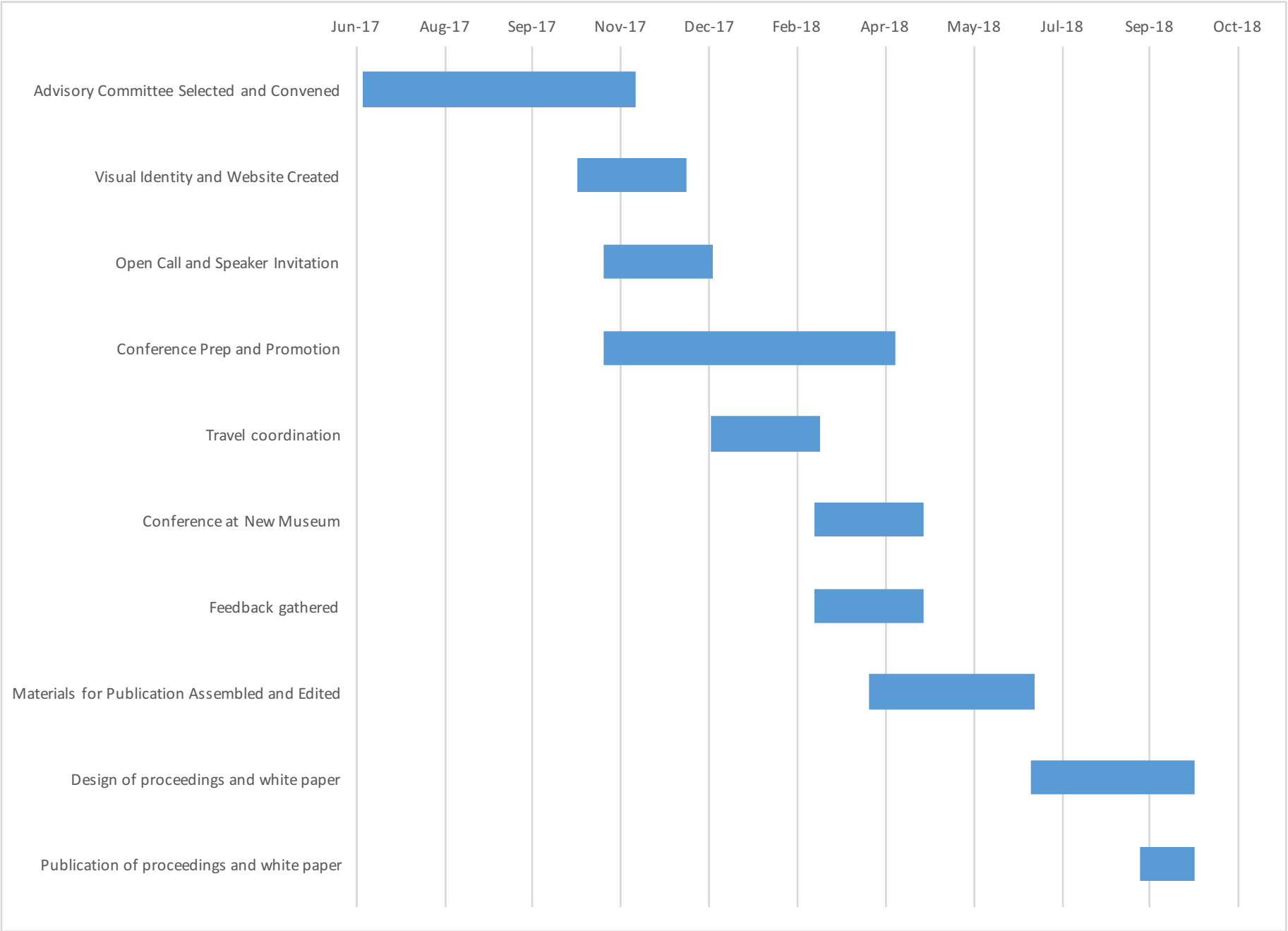
- My organization is better prepared to provide a program or service that addresses community needs.
- My organization is better able to engage my community.
- My organization is better prepared to develop and maintain on-going relationships with community partners.
- My organization is better prepared to share knowledge and other resources as an active contributor to problem solving in the community.

Our secondary aim would be for 50% of artists, activists, and community representatives in attendance to answer affirmatively to one of these questions:

- This conference and white paper addressed community needs with regard to web archiving.
- The program actively contributed to problem solving in the community.

*What products will result from this project?*

The project will result in a live-streamed three-day conference, a website featuring video documentation of the project, a published proceedings, and a white paper offering a set of principles and best practices on ethics in web archiving



## DIGITAL PRODUCT FORM

### Introduction

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (i.e., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. However, applying these principles to the development and management of digital products can be challenging. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

### Instructions

You must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

## PART I: Intellectual Property Rights and Permissions

**A.1** What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

The white paper produced will be published under a Creative Commons license. Contributors to the proceedings will be asked to select a license for their work appropriate to their situation, but will be encouraged to use a Creative Commons license.

**A.2** What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

Rhizome will impose no further conditions, apart from asking that rights asserted by conference participants be respected. Regardless of license, all conference materials will be free to access on the conference microsite.

**A.3** If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

Conference participants will be asked to sign clearance forms for their participation and the potential publication of their words. A conference code of conduct will be developed by the organizing team to ensure that participants are careful in their handling of any potentially sensitive material or relationships.

## Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

### A. Creating or Collecting New Digital Content, Resources, or Assets

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

We will create a public website with information about the three-day conference, a call for participation, conference livestream and video documentation, published proceedings (20,000 words, 20 illustrations, PDF, also available as print-on-demand), and white paper (two pages, PDF) We will collect applications to participate in the conference (private, held in Rhizome's organizational archives).

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

The website will be set up using a content management system such as Drupal with a MySQL database. The PDF will be designed using Adobe InDesign and image editing will be done in Adobe Photoshop using MacBook Pro computers.

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

Videos will be archived as .MP4, 1920x1080, H.264, AAC. Images will be collected as uncompressed TIFF or PNG where possible. PDFs will be produced at 300DPI resolution for print-on-demand and 72DPI resolution for online circulation.

## **B. Workflow and Asset Maintenance/Preservation**

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

Conference planning will be overseen by the co-organizers via a weekly planning meeting, with monthly reports to the advisory group to monitor progress against predetermined milestones. Following the conference, Rhizome's Editorial Manager will check in weekly with the co-organizers to assess progress and quality of work.

**B.2** Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and 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 C.F.R. § 200.461).

Preservation of the assets created as part of this conference will be included in the Rhizome Artbase and the New Museum digital archive.

## **C. Metadata**

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

Conference materials will be logged in Rhizome's Artbase database using a metadata structure based on Dublin Core. Metadata content will be drawn from The Art & Architecture Thesaurus (AAT) and

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

Metadata will be stored in Rhizome's Artbase, which runs on Wikibase. It will also be included in the New Museum's Digital Archive, which runs on Collective Access.

**C.3** Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

Rhizome's Wikibase implementation is an early-stage effort to develop a digital archive database that can facilitate widespread information sharing.

## **D. Access and Use**

**D.1** Describe how you will make the digital content, resources, or assets 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).

The digital content produced during this project will be made available via the project website for download, openly available online, from Rhizome's web server infrastructure, which is located across Amazon Web Services and Google Cloud. Live streaming will be done via the New Museum's subscription at livestream.com. Videos will be made accessible via Rhizome's Pro account on Vimeo.com.

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

Rhizome Artbase - <https://rhizome.org/art/artbase/>

Webrecorder - <http://webrecorder.io>

## **Part III. Projects Developing Software**

### **A. General Information**

**A.1** Describe the software you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) it will serve.

**A.2** List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

### **B. Technical Information**

**B.1** List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

**B.2** Describe how the software you intend to create will extend or interoperate with relevant existing software.

**B.3** Describe any underlying additional software or system dependencies necessary to run the software you intend to create.

**B.4** Describe the processes you will use for development, documentation, and for maintaining and updating documentation for users of the software.

**B.5** Provide the name(s) and URL(s) for examples of any previous software your organization has created.

### **C. Access and Use**

**C.1** We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

**C.2** Describe how you will make the software and source code available to the public and/or its intended users.

**C.3** Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository:

URL:

### **Part IV: Projects Creating Datasets**

**A.1** Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

**A.2** Does the proposed data collection or research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity been approved? If not, what is your plan for securing approval?



**A.3** Will you collect any personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information? If so, detail the specific steps you will take to protect such information while you prepare the data files for public release (e.g., data anonymization, data suppression PII, or synthetic data).

**A.4** If you will collect additional documentation, such as consent agreements, along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

**A.5** What methods will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

**A.6** What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

**A.7** What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

**A.8** Identify where you will deposit the dataset(s):

Name of repository:

URL:

**A.9** When and how frequently will you review this data management plan? How will the implementation be monitored?