## Board of Trustees of the University of Illinois

# Data Storytelling Toolkit for Librarians

## Brief Abstract:

The Data Storytelling Toolkit for Librarians (DSTL) will connect real-world examples of data use with data stories (including narrative strategies and data visualizations) as adaptable templates for communicating and related online free resources for data exploration and visualization.

## **Statement of Broad Need**

Greater datafication in all aspects of our lives presents challenges and opportunities for storytelling as a powerful way of conveying meaning. Public libraries (PLs) and community college libraries (CCLs) face related challenges of communicating their offerings and impacts to the public, which storytelling can address. Data storytelling means any presentation of data that uses narrative strategies, and "data stories" here mean collections of resources that provide templates for stories about common library data and why it matters. A story can show the audience what matters about data and help them understand its relevance. Through data storytelling, library staff can make data and information into meaningful and powerful communication, for themselves and their communities.

Earlier library data collection models focused on measuring library work through quantifiable outputs, reducing services to numbers that often did not provide useful measures of effectiveness. (Dresang, 2006) A field-wide shift occurred with the 2013 Public Library Association launch of Project Outcome, closely followed by the IMLS-funded 2015 Research Institute for Public Libraries (RIPL). Both projects retool the data collection process to focus on measuring the outcomes achieved by library work. This shifted data practices from collecting numbers that demonstrate "who/what/when" of library services, to collecting data that can be used to show "how" libraries impact their communities and "why" libraries matter. The launch of the Public Library Data Alliance (PLDA) in July 2020 signals another stage in making data meaningful in library conversations and cultural expectations. 2020-21 RIPL Data Boot Camp events regularly see over 200 participants per session (https://www.ripleffect.org/ripl-data-boot-camp/). These IMLS-funded projects have contributed greatly to understanding the importance of data collection in the library field and instigating a culture shift, supporting librarians in the analysis, interpretation, and promotion of the data they collect. These projects have blazed a much-needed trail for better evaluation of library activities and services.

But there is still work to be done. These tools help librarians learn how to collect better library data with the expectation of engaging stakeholders and building library advocates. However, they do not provide guidance in how to design data stories that will make data meaningful, memorable, and center data at the heart of library stories. Rather than being opposites, data and storytelling can coexist in a powerful partnership between data outcomes, storytellers, and audiences. And, despite its 120+ year history in youth services, most librarians still overlook the well-documented power of storytelling and stories for captivating audiences and making messages both memorable and retellable. (K. McDowell, 2018, 2020; K. and E. M. C. McDowell, 2017) At the University of Illinois, we have been bringing these ideas together in the classroom for the past four years. Our work has shown the power of combining the fields of data analysis and storytelling in the information sciences. Now is the time when libraries need a

data storytelling toolkit, to increase powerful communication through the pressures of economic instability and COVID-19.

The goal of this project is to plan, develop, and pilot a Data Storytelling Toolkit for Libraries (DSTL), that builds on current data collection best practices to create a toolkit to train librarians in meaningful and powerful communication using their collected data through data storytelling. Data storytelling is an emerging area, and we offered the first data storytelling course (to the authors' knowledge) at the School of Information Sciences at Illinois. The course is co-taught by PI Kate McDowell who has taught library and information science storytelling and led a storytelling festival for 13 years and co-PI Matt Turk who directs the Data Exploration Lab and brings data visualization expertise from across information science and astronomy. This course content has great potential for transforming the library field by helping library staff to tell more effective, impactful stories. However, the syllabus will need transformation to effectively serve as a toolkit for libraries 1) to meet the needs of professional librarians and their obstacles to data storytelling, 2) for efficient delivery in workshops, and 3) in a digitally oriented pandemic and post-pandemic context.

We have begun to investigate translation from a student-focused audience to library venues for data storytelling concepts. Thanks to a 2020 seed grant from the University of Illinois Center for Social & Behavioral Science, 26 librarian interviews have gone through thematic analysis (Braun, Clarke, Hayfield, & Terry, 2019) and a preliminary Data Storytelling Obstacles Questionnaire has received 38 responses. Interviewees and questionnaire respondents came from iSchool alumni (in PLs and academic libraries including CCLs), the Illinois Library Association (ILA, https://www.ila.org/), the Consortium of Academic and Research Libraries in Illinois (CARLI, with high CCL membership, https://www.carli.illinois.edu/), and the national IMLS-funded project Measures that Matter (MtM, https://measuresthatmatter.net/).

Initial interview findings include that libraries are predominately retelling the same stories they have told for a long time, regardless of their impact on a specific audience or effectiveness at building library support. Initial questionnaire responses indicate that, of 8 possible obstacles librarians experience when attempting to tell data stories, the top three are: too little time; don't have data they need; and too little focus on storytelling with not enough connection to data. Respondents, including some participants already familiar with Project Outcome, RIPL, and the new Public Library Data Alliance (launched July 2020 as the next steps for MtM), have voiced a desire for clear information to teach them to better apply the data they collect in their work. Interview comments emphasize that even when the "right" type of data is being collected, librarians are unclear on how to use that data to support their library's strategic mission, promote their work, engage stakeholders, or advance the scope of their practices. If they are familiar with data storytelling at all, they lack confidence in how to design effective data stories to support their work.

We see an enormous gap where data storytelling could serve to make both librarian data usage and their stories of practice more current and compelling, prompting more reflective data collection practices that are directly connected to communication goals. We propose to collect data use scenarios from PLs and CCLs, beginning from an understanding that there are some commonalities in the ways that data must be used to explain, justify, or advocate for organizational priorities across libraries. This project will connect data use scenarios with data story tools, using journey mapping (actions, motivations, questions, barriers) to unpack the data life cycle in practical action for libraries. Data stories will connect each data use scenario with a template for building a story and communicating it via storytelling, including journey map,

narrative strategies, and data visualizations. The DSTL will be based on the most common data use scenarios, redesigned as solution-focused data stories (data use scenario with journey map, narrative strategies, and data visualizations) for each scenario and related online free resources for data exploration and visualization (including Twine <u>https://twinery.org/</u>; Vega-Lite <u>https://vega.github.io/vega-lite/</u>; Observable <u>https://observablehq.com/demo</u>; GitHub <u>https://github.com/</u> and more).

The Internal Advisory Team (IAT) will include: PI Dr. Kate McDowell, storytelling expert; co-PI Dr. Matthew Turk, data visualization expert; Dr. Martin Wolske, expert in community informatics; Dr. Rachel Magee who empowers librarians to conduct research; Dr. Kyungwon Koh, expert in the library maker movement; and Dr. Sharon Comstock who serves on the Public Library Data Alliance board. The IAT will be consulted on an ongoing ad hoc basis, and especially as the project reaches evaluation milestones. Each IAT member's expertise will be important as the project moves through each of the three phases, providing fresh insights from a place of familiarity with the project.

This project meets the "build capacity" goal and the "share and adopt best practices" objective, and supports all other categories of objectives within that goal. In addition, departments that provide library services to under-represented communities are put at a disadvantage when libraries rely on the traditional library stories told and data collected, as these often favor the needs of the communities already receiving the majority of services. Through partnership with libraries with strong Diversity, Equity, and Inclusion infrastructures (DEI) we aim to close some of these equity gaps.

Frankly, there are thousands of librarians who are skittish about data but love stories. And there are hundreds of librarians who see data as fundamental, but until those librarians have a language through which to connect with the passions of the thousands who love stories, this movement toward strategic data use in the field of libraries will be stifled, along with the potential collaborative creativity of librarians. This toolkit will provide a set of easy-to-adapt templates, meant to inspire creative engagement with data even by those who remain skeptical. In the best traditions of folklore, the toolkit will be clear enough to help librarians see the potential for communicating with data, but flexible enough to be fully adaptable to each librarians' setting and to communication needs inside the organization and with the public. As in folklore, the best stories are told again and again, and finding compelling library data story archetypes could energize our field to embrace data with story in ways that might unleash the best creativity of our own practitioners.

Compelling data stories can be told again and again, but they must engage the audiences they address. Similarly, it will be important that the workshops be responsive to participants, in the moment and over time. Every workshop will involve active learning tailored to the online situation, highlighting collective librarian interests in order to motivate creative interactions in real time. Workshops will be recorded, and so each one should reinforce previous ideas while also bringing something vital and timely from the moment of its presentation. The overall goal will be to ensure that, while some ideas are reinforced, there is a consistently inspiring aspect to each workshop. McDowell won a major campus-level award for her track record as online instructor since 1998, and these workshops will draw from those strengths.

Librarians know the importance of their work, but building confidence in data storytelling empowers them to communicate better and build support with essential stakeholders. Data storytelling is already appearing as a best practice within industries outside of the library profession that libraries compete with for needed political support and financial grants. To stay relevant within these contexts, librarians must develop a stronger understanding of data storytelling practices, and the DSTL will enable librarians to do so. Project supporters include RIPL nationally, CARLI connecting academic libraries in Illinois, NILRC connecting community college libraries in Illinois, CCL connecting community college libraries in California, and several other PLs and CCLs (see support letters).

## **Project Design**

Creating the DSTL requires three phases. In Phase 1 (six months), we will collect data use scenarios via surveys, journey maps, interviews, and a workshop with Core Design Team (CDT) comprised of professional librarians who regularly communicate library data to public stakeholders. In Phase 2 (twelve months), we will create data stories as the primary content of the DSTL and evaluate and improve these data stories through a workshop-based collaborative iterative design process. In Phase 3 (six months), we will pilot the data storytelling toolkit and seek final refinements via public input from librarian audiences. The CDT will reconvene for final input after the pilot sessions.

## Phase 1 (six months, 9/21-2/22)

The goal of this phase is to **collect data use scenarios** (such as library impact on employment outcomes for community members) as the basis for the toolkit design.

A data story to communicate library impact on employment will include story and image strategies for **who** has become employed based on library services, **how** (journey map showing a visual sequence of steps from job seeking to employment), a structure for the **story** of an individual's outcomes, and a strong **data visualization** strategy for communicating this impact.

This phase will require recruiting participating libraries and librarians and assembling a Core Design Team (CDT) in order to rigorously evaluate the relevance of data use scenarios in order to create a powerful set of examples. For three months, the PIs and research assistant (RA) will recruit participants for the CDT, consisting of invited professionals from libraries with strong DEI commitments whose work directly involves communicating library data to stakeholders, both within the organization and publicly. At least one person from each participating library will be invited. CDT members will complete an initial survey to gather data about prior data storytelling successes and challenges. For the next three months, PIs will conduct design interviews with CDT members, using journey maps that will visually capture data flow, from collection to storytelling and other reporting. (Richardson, 2010) The online digital tool Miro (https://miro.com/) is a good example of how journey maps can be captured online at a distance. Journey maps will help the team understand how data is currently communicated to various levels of library stakeholders (between departments, from/to administrators/boards, from/to the public).

PIs and the RA will synthesize findings and present the most prevalent data use scenarios to the CDT. The team will collaborate via online workshops to refine generalizable descriptions of at least five common scenarios. The result will be a set of data use scenarios that will form the basis for the toolkit in Phase 2, and a list of related online free resources for data exploration and visualization to support toolkit use. The CDT will also be asked to complete two surveys (preand post-workshop). The CDT and The Internal Advisory Team (IAT) will be asked to give feedback on a semi-structured focus group script for use in Phase 2. Data collection for this phase will include interviews, journey maps, and surveys; data analysis will be ongoing throughout the project.

## Phase 2 (twelve months, 3/22-2/23)

The goal of this phase is to **create data stories as the content of the DSTL**. Data stories will connect each data use scenario with a template for storytelling, including typical journey map, narrative strategies, and data visualizations.

A data story about library employment impacts might include a **journey map** that visualizes the flow of data from statistics on computer usage to survey of computer lab users to presentation to board of trustees. **Narrative strategies** might include the hero's journey (Campbell, 1949) as a success story of a job seeker as character, or enigma code (Barthes, 1974) to follow the suspense from mystery to investigation to outcome. **Data visualization** might show aggregate changes in demand for employment support at libraries over time.

During the first three months, PIs with RA support will design data stories that include typical journey maps, useful narrative strategies (such as identifying a "character" or narrative structure), and data visualizations. During the next six months, a collaborative iterative design process will be used in order to rigorously evaluate the relevance and efficacy of the tools proposed for each scenario. For the next three months, design workshops (one per month) will invite participants to collaboratively improve data stories (data use scenario with journey map, narrative strategies, and data visualizations). Participants will test the data stories via storytelling and audience feedback in online workshops (with both large group and break-out group sessions) and post-workshop surveys for further data story evaluation and workshop assessment. . PIs and RA will conduct focus groups and interviews to further test and refine the data stories as storytelling templates.

All participating libraries and CDT members will be invited to these interactive workshops and to at least one focus group using the semi-structured script developed in Phase 1. Interviewees will be selected based on expertise. PIs will continue to add to the list of free online tools to support toolkit use. The CDT will review the data stories and the draft toolkit prior to release of the pilot toolkit. Data collection for this phase will include workshops, workshop evaluation surveys, focus group transcripts, and follow-up interview transcripts.

## Phase 3 (six months, 3/23-8/23):

The goal of this phase is to pilot the data storytelling toolkit for final refinements via public input from librarian audiences. Through a series of five monthly public workshops, the PIs and RA will test the data stories (data use scenarios with journey map, narrative strategies, and data visualizations for each) and solicit audience feedback via post-workshop evaluation surveys.

> The data story about library employment impacts might need to change to show more steps in the **journey map**, including data about how employment services are marketed to the public. An additional **data visualization** might be added to show evidence of community need for employment support services.

These public workshops will be widely advertised (to groups such as RIPL alumni) and followed by post-workshop surveys for additional input. One final workshop with the CDT and IAT will be used to finalize the toolkit and list of related online free resources to support toolkit use in specific contexts. Findings from the data analysis will be prepared for publication and conference presentation (PLA, ALA, ACRL, etc.).

Phase	Activities	9/21- 11/21	12/21- 2/22	3/22- 5/22	6/22- 8/22	9/22- 11/22	12/22- 2/23	3/23- 5/26	6/23- 8/23
Phase 1	Recruit Core Design Team (CDT)								
	Survey								
	Design interviews using journey maps								
	Data use scenarios collected and refined								
	CDT virtual workshop								
	CDT surveys, CDT and IAT feedback								
Phase 2	Create data stories for each use scenario								
	Design workshops (one per month)								
	Data stories tested and refined								
Phase 3	Pilot DSTL workshops								
	Final CDT and IAM workshop								

Performance Measurement	Assessment plan and data collected
Outputs	<ul> <li>Phase 1: Assessed by successful collection of data use scenarios.</li> <li>Phase 2: Assessed by whether the project yields a workable DSTL design for phase 3, CDT workshops and post-workshop surveys.</li> <li>Phase 3: DSTL public presentations assessed via post-workshop surveys.</li> </ul>
Outcomes	Phase 1: Assessed by pre- and post-workshop surveys Phase 3: Assessed by post-workshop surveys
Efficiency	Phase 3: Assessed by post-workshop surveys
Quality	Phase 2: Assessed by post-workshop surveys, interviews, focus groups Phase 3: Assessed by post-workshop surveys, interviews, focus groups
Timeliness	Assessed by co-PIs, at 3-month intervals

Performance measurement data will be collected in addition to the data for required performance reports. The project plan is designed in 3-month intervals to support timely completion.

## **Diversity Plan**

Diversity is foundational to storytelling in listening, meaning that storytelling only occurs where there is an audience willing to listen. As a field with diversity challenges, the design of this project is meant to focus on listening to librarians. As we recruit, however, we will be thoughtful about soliciting participation from libraries that have worked to increase diversity, in their services and in the field. The first stage will focus on recruiting Core Design Team (CDT) members from libraries with a history of strong diversity programs and then requesting their input as to other diversity leaders to invite. Libraries will be identified based on their reputation as leaders in the work of Diversity, Equity, and Inclusion (DEI) in the library field. These libraries will share a focus on reaching underserved populations and have demonstrated their DEI leadership through their departments, librarian positions, or task forces designed to specifically for to support DEI initiatives; by their investment in the state of DEI in the field by leading trainings at national and local-level conferences; and for their innovative approaches to patron service models that creatively address the needs of under-served communities. Selection of a preliminary group has already been made by focusing on how DEI is reflected on library websites, but we will ask CDT members to invite libraries with strong track records who may or may not represent their accomplishments on their websites. Invited libraries will be geographically, demographically, and socio-economically diverse, representing a wide range of regional locations, service population sizes, and community identities.

#### Board of Trustees of the University of Illinois

In order to make the resulting data storytelling toolkit as relevant as possible to a diverse audience, DSTL will design for inclusion by incorporating approaches from critical race theory and research on library services to diverse populations in the toolkit and subsequent workshops. (Cooke, 2016; Gibson & Hughes-Hassell, 2017) This will be a foundational value in all areas of DSTL including the project framing, the collection of data use stories, and the development of the pilot toolkit and its data storytelling strategies.

#### **Broad Impact**

Data storytelling was first taught in our iSchool, but it is now beginning to be taught in journalism and business. Libraries must keep up with these transformations in communication expectations as we compete for funding. Savvy storytelling is already key for libraries to sustain public knowledge of their accomplishments, but data storytelling can position libraries to make gains in community buy-in and financial support. Further, even practitioners who are reluctant to engage with data can immediately see the value of storytelling, and so the connection of data and story in the DSTL has great for buy-in from librarians of all kinds, ultimately transforming attitudes and practices across librarianship.

The goal of this planning grant is to develop a toolkit, based on the emerging area of data storytelling, that will ultimately support effective data-driven communication in libraries. This toolkit will empower public and academic librarians to become data storytelling practitioners. The DSTL will build on the significant gains made by other IMLS data-focused projects. Other projects engage data but not storytelling, and this project can help bridge that gap. Project Outcome focuses on collecting and sharing outcomes, but not connecting those data to the bigger picture stories libraries tell. MtM and the PLDA focus on data communication, but as a subsidiary of their primary mission to boost data conversations in the library world. A recent RIPL communication workshop presented data as opposed to story. No other library-centered projects yet focus on building stories from data. This proposal sees allies in RIPL, Project Outcome, MtM and PLDA in order to build on these cultural shifts in the library community toward demonstrating the value of data and cohesive data collection practices. The MtM mailing list and RIPL alumni list have already been made available to the PIs as venues for recruiting Core Design Team (CDT) members and advertising public workshops, and we expect ongoing dialogues with PLDA via project team member and PLDA board member Dr. Sharon Comstock.

Storytelling is one of the most powerful human communication forms and has a rich history in the field of library and information science. Data storytelling is rapidly growing to meet the demands of workforces increasingly saturated with data. This proposal builds on the iSchool's foundational work in this field and aims to bring these powerful tools to practicing libraries. As these practitioners are often the community portals to data and information, they are vital in bridging gaps between community members and our increasingly datafied world. Empowering community libraries with data storytelling tools empowers communities to effectively communicate information to a diverse set of stakeholders. The DSTL will empower libraries "to spark rapid change in their communities"<sup>1</sup> with memorable data-focused stories. The DSTL will extend a new focus on data in libraries, providing a library-specific toolkit for encouraging purposeful and memorable data communication through storytelling. In future work, the DSTL will support the long-term goal to create a network of librarians as data storytelling

<sup>&</sup>lt;sup>1</sup> <u>https://www.imls.gov/blog/2018/03/2018-community-catalysts-funding-opportunity-chance-innovate</u>

#### Board of Trustees of the University of Illinois

experts serving communities nationally. This planning project could lead to a future proposal with "train the trainer" model distribution, so that these planning-project participants may ultimately carry the toolkit forward into all of their professional networks and communities. The long-term vision is to transform librarianship not only by centering data storytelling as organizational communication practice, but also by cultivating data storytelling expertise as a signature expertise of our field, so that when communities have data storytelling needs, they are met at libraries and by librarians.

# Schedule of Completion

Phase	Activities	9/21- 11/21	12/21- 2/22	3/22- 5/22	6/22- 8/22	9/22- 11/22	12/22- 2/23	3/23- 5/26	6/23- 8/23
Phase 1	Recruit Core Design Team								
	Survey								
	Design interviews using journey maps								
	Data use scenarios collected and refined								
	Core Design Team virtual workshop								
	Core Design Team workshop surveys								
Phase 2	Create data stories for each use scenario								
	Design workshops (one per month)								
	Data stories tested and refined								
Phase 3	Pilot DSTL workshops								
	Final Core Design Team workshop to								



## DIGITAL PRODUCT FORM

#### INTRODUCTION

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to digital products that are created using federal funds. This includes (1) digitized and born-digital content, resources, or assets; (2) software; and (3) research data (see below for more specific examples). Excluded are preliminary analyses, drafts of papers, plans for future research, peer-review assessments, and communications with colleagues.

The digital products you create with IMLS funding require effective stewardship to protect and enhance their value, and they should be freely and readily available for use and reuse by libraries, archives, museums, and the public. 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

If you propose to create digital products in the course of your IMLS-funded project, you must first provide answers to the questions in **SECTION I: INTELLECTUAL PROPERTY RIGHTS AND PERMISSIONS.** Then consider which of the following types of digital products you will create in your project, and complete each section of the form that is applicable.

#### SECTION II: DIGITAL CONTENT, RESOURCES, OR ASSETS

Complete this section if your project will create digital content, resources, or assets. These include both digitized and born-digital products created by individuals, project teams, or through community gatherings during your project. Examples include, but are not limited to, still images, audio files, moving images, microfilm, object inventories, object catalogs, artworks, books, posters, curricula, field books, maps, notebooks, scientific labels, metadata schema, charts, tables, drawings, workflows, and teacher toolkits. Your project may involve making these materials available through public or access-controlled websites, kiosks, or live or recorded programs.

#### SECTION III: SOFTWARE

Complete this section if your project will create software, including any source code, algorithms, applications, and digital tools plus the accompanying documentation created by you during your project.

#### SECTION IV: RESEARCH DATA

Complete this section if your project will create research data, including recorded factual information and supporting documentation, commonly accepted as relevant to validating research findings and to supporting scholarly publications.

#### SECTION I: INTELLECTUAL PROPERTY RIGHTS AND PERMISSIONS

**A.1** We expect applicants seeking federal funds for developing or creating digital products to release these files under open-source licenses to maximize access and promote reuse. What will be the intellectual property status of the digital products (i.e., digital content, resources, or assets; software; research data) you intend to create? What ownership rights will your organization assert over the files you intend to create, and what conditions will you impose on their access and use? Who will hold the copyright(s)? Explain and justify your licensing selections. Identify and explain the license under which you will release the files (e.g., a non-restrictive license such as BSD, GNU, MIT, Creative Commons licenses; RightsStatements.org statements). 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.

**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.

**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.

#### SECTION II: 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 the format(s) you will use.

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

**A.3** List all the digital file formats (e.g., XML, TIFF, MPEG, OBJ, DOC, PDF) you plan to use. If digitizing content, describe the quality standards (e.g., resolution, sampling rate, pixel dimensions) you will use for the files you will create.

#### Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan. How will you monitor and evaluate your workflow and products?

**B.2** Describe your plan for preserving and maintaining digital assets during and after the award period. Your plan should 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).

#### Metadata

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

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

**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).

#### 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, delivery enabled by IIIF specifications).

**D.2**. Provide the name(s) and URL(s) (Universal Resource Locator), DOI (Digital Object Identifier), or other persistent identifier for any examples of previous digital content, resources, or assets your organization has created.

#### SECTION III: SOFTWARE

#### **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 or similar functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

#### **Technical Information**

**B.1** List the programming languages, platforms, frameworks, 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), URL(s), and/or code repository locations for examples of any previous software your organization has created.

#### Access and Use

**C.1** Describe how you will make the software and source code available to the public and/or its intended users.

**C.2** Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository:

URL:

#### SECTION IV: RESEARCH DATA

As part of the federal government's commitment to increase access to federally funded research data, Section IV represents the Data Management Plan (DMP) for research proposals and should reflect data management, dissemination, and preservation best practices in the applicant's area of research appropriate to the data that the project will generate.

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

**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 sensitive information? This may include personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information. If so, detail the specific steps you will take to protect the information while you prepare it for public release (e.g., anonymizing individual identifiers, data aggregation). If the data will not be released publicly, explain why the data cannot be shared due to the protection of privacy, confidentiality, security, intellectual property, and other rights or requirements.

**A.4** What technical (hardware and/or software) requirements or dependencies would be necessary for understanding retrieving, displaying, processing, or otherwise reusing the data?

**A.5** What documentation (e.g., consent agreements, data documentation, codebooks, metadata, and analytical and procedural information) will you capture or create along with the data? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the data it describes to enable future reuse?

**A.6** What is your plan for managing, disseminating, and preserving data after the completion of the award-funded project?

**A.7** Identify where you will deposit the data:

Name of repository:

URL:

**A.8** When and how frequently will you review this data management plan? How will the implementation be monitored?