Fedora Migration Paths and Tools: A Pilot Project

The DuraSpace Community Supported Programs division at LYRASIS proposes an 18-month project to develop, pilot, and document tools and paths to migrate from the Fedora 3 repository software, which is widely used but no longer supported, to Fedora 6. The proposed project is the result of a year-long planning effort funded by IMLS National Leadership Grant LG-72-18-0204-18, which investigated barriers to migrations and upgrades at institutions running older versions of Fedora. Hundreds of libraries and archives in the U.S. use Fedora repository software to deliver scholarly, scientific, and cultural heritage resources to patrons. Continued reliance on Fedora 3 puts the stability, security, accessibility, and functionality of these repositories at risk. Through two in-depth pilots, LYRASIS and the Fedora community will produce tools, documentation, paths, case studies, and best practices to support and improve the migration process for Fedora 3 repositories. Providing support for Fedora 3 migrations will encourage implementation of an up-to-date system in Fedora 6, one that is more sustainable, expands repository functionality, and supports long-term content preservation. The pilot project supports IMLS's goal to "Build Capacity" by strengthening the ability of libraries and archives to sustain repositories that are critical to providing access to digital content and collections.

Statement of National Need

Fedora (https://duraspace.org/fedora/) is a flexible, modular, open source repository system for the management and dissemination of content through digital libraries and archives, including specialized access to large and complex digital collections of historic and cultural materials as well as scientific data. Fedora's architecture is built on the principle that interoperability and extensibility is best achieved by providing a limited set of stable, standards-based repository services and common patterns for integrating with other best-practice systems and applications. These services are provided via RESTful APIs. Fedora offers a foundation upon which many types of repository frameworks can be built, including the Islandora (https://islandora.ca/) and Samvera (https://samvera.org/) repositories.

Fedora was first released in 2003. The currently supported versions, 4 and 5, were released in 2015 and 2018, respectively. Fedora 6 is now under development with public release anticipated in early 2021. However, more than two-thirds of Fedora installations (269 institutions) are still using Fedora 3, which was released in 2008, and is no longer being supported. One of the challenges for Fedora 3 institutions is that Fedora 4 is a completely re-written application, designed to address performance and scale issues, that shares no code with Fedora 3. By contrast, Fedora 5 and 6 are built on the same codebase as Fedora 4; their version numbers correspond with the adoption of semantic versioning (https://semver.org) best practices. The differences between Fedora 3 and Fedora 4, 5, or 6 impact underlying technologies, data models, standards for description, and functionality, meaning that an upgrade requires both re-modeling and migrating data, making the move time-consuming, costly, and challenging to support.[1] Fedora 4 and 5 share underlying technologies with version 6; an upgrade will still require effort, but significantly less so than for Fedora 3. In addition, the release of Fedora 6 will include a migration utility that will provide a streamlined upgrade path for Fedora 4 and 5 implementations. With the majority of Fedora installations still dependent on version 3, and the challenges inherent in upgrading it, Fedora's leadership decided to focus this pilot migration project on moving from Fedora 3 to 6. Fedora 4 and 5 users will still benefit from some of the processes documented through this project, such as data modeling and testing procedures.

Most Fedora 3 repositories emphasize digital preservation and library-managed, at-risk collections. A single repository can store millions of objects and hundreds of terabytes of unique content, providing access

---

[1] Within the context of this proposal, "upgrade" generally refers to a situation where a new application is dropped on top of data with relatively little effort; in other words, the data does not need to be exported from the old application and imported into the new one. A migration involves taking data from one application (for example, Fedora 3), changing it to comply with a new system, and importing it into the new application (Fedora 4, 5 or 6).

to a wide array of resources that support teaching, research, and public engagement. Among the 111 respondents to a 2019 survey of Fedora users conducted as part of the preceding planning grant (IMLS LC-72-18-0204-18), the most common forms of content stored in their repositories were cultural heritage and special collections (86%), archives (78%), scholarly publications (57%), and research data (47%). One-third of respondents estimated that their repositories hold more than 500,000 digital objects; 26% hold 100,000 to 500,000 objects. Fedora 3 repositories include digital books, manuscripts, maps, photographs, oral histories, music, video, data, web sites, theses, dissertations, journal articles, 3D objects, and other formats. The intellectual output of the campus and collections of scholarly and educational research materials are stored in an organization's Fedora 3 repository, which can also support open access.

Fedora 3 is used by a variety of sizes and types of organizations including liberal arts colleges, library consortia, special libraries, historical organizations, and moderate or higher research activity universities.[2] About one-third of Fedora 3 repositories are highest activity (R1) research institutions within the Carnegie Classification® system.

- The majority of the Islandora community, including 163 repositories in the U.S., is running Fedora 3, which underlies Islandora 7 and earlier versions. The recently released Islandora 8 integrates with Fedora 5 (and will with Fedora 6). Only 12% of the Islandora repositories in the U.S. are hosted by R1 universities. The Islandora community is generally characterized by smaller, less-resourced institutions and organizations, including Carnegie-classified R2 and R3 universities such as the University of Denver and Andrews University, liberal arts colleges such as Vassar College and Barnard College, library consortia such as the Metropolitan New York Library Council, and historical societies such as the Adventist Digital Library and the California Historical Society.
- Nearly 90 custom repositories are built on Fedora 3 in the U.S., 48% of which are hosted by R1 universities. In addition, there are 20 Samvera repositories using Fedora 3 in the U.S., 85% of which are hosted by R1 universities. Using Geoffrey Moore's famous Technology Adoption Cycle, Fedora custom and Samvera implementers are more likely to fall into the innovator or early adopter categories, willing to adopt new technologies and buy into new product concepts very early in their life cycle. They also tend to be better resourced and have more technology staffing. Members of the Samvera and Fedora communities that are not R1 institutions include Amherst College, Lafayette College, and Science History Institute (formerly the Chemical Heritage Foundation).

Interviews done by DuraSpace in 2017 with 31 community members gathered user stories about major upgrades and migrations, eight of which came from pioneers in the upgrade path from Fedora 3 to 4. Their stories showed that the challenges faced during migration were significant, including issues with metadata normalization, gaps in skills and knowledge, keeping pace with development in the repository communities, and redefining services based on the new capabilities of Fedora 4. One example that illustrates this came from an interviewee who experienced difficulty with two of Fedora 4's dependencies for their particular use cases. The interviewee said the migration "failed because at the time we were the largest repository to try to migrate…. There were problems with the underlying Infinispan and ModeShape components when dealing with large files." This story led to a community discussion about how Fedora 4 provides preservation support. The interviewee said: "When we move to F[edora]4, we'll have to totally change the way we do preservation."

To better understand issues and community needs, a year-long planning effort was funded by IMLS National Leadership Grant LG-72-18-0204-18 to explore the barriers to migrating Fedora 3 repositories and recommend strategies to lessen these. The assessment included an environmental scan, review of sample Fedora 3 institutional profiles, identification and review of relevant technologies that could support migration,

---

[2] An Internal analysis of Fedora 3 institutions and repository platforms is available at https://docs.google.com/spreadsheets/d/1xIrB8EJqFRWlRtmdwTEoT2xMSGxE7MMryZGbD3eqUs8/edit#gid=0.

a survey of the Fedora community, and a focus group with 11 participants to explore survey results. Findings are documented in a Final Report included in Supportingdoc5 and available at https://wiki.lyrasis.org/display/FF/Designing+a+Migration+Path+-+Final+Report.

Motivating factors for repository migrations vary by institution and could include such factors as cost reduction, improvements in scale and performance, expanded services, and better integration to support related applications like digital preservation. The factors that make migration difficult also vary by organization and repository, but the most frequently cited have to do with metadata – inconsistent data, custom metadata fields, different data models between repositories, and the need for metadata mapping. Several case studies that were reviewed reported significant time spent on metadata normalization, deduplication, and remediation.[3] A related difficulty involves inconsistent or "messy" source data. The process of mapping metadata from one repository system to another would be much simpler were it not for metadata quality problems in legacy systems like custom local fields, duplicate fields, and misspelled entries.

The planning grant's survey of the Fedora community provided more insight into specific challenges faced when moving from one version of Fedora to another. Among the 111 respondents, 79% of whom are using Fedora 3, the most pressing barrier to upgrading repository software was availability of staff and funding (68%). Among respondents who had attempted an upgrade from Fedora 3 to version 4 or 5, lack of resources was also cited as a barrier by 49%. While the time and cost of moving from Fedora 3 to later versions vary widely based on multiple factors (repository size, content, data models, metadata quality, standards, etc.), anything that can reduce the burden on local staff will help facilitate upgrades and migrations. Most survey respondents (45%) reported having 1-3 full-time employees (FTE) working on their repository; 12% reported less than one FTE to manage their repository. In the Bridge2Hyku project, partners reported that only one-third of the staff involved in digital collections management were from IT, an average of 2.5 FTE among their institutions.[4] An upgrade and migration are additional burdens on staff already occupied with managing the existing digital content and repository system. Other factors that served as barriers to migration for those who upgraded (or attempted to) from Fedora 3 to 4 or 5 were lack of compatibility with front-end applications (45%), changes in metadata standards (40%), and issues with performance and scale (33%). When asked what would help in upgrading, the top three survey responses were content migration tools (71%), metadata migration tools (61%), and documentation (52%).

The planning grant also explored Fedora and eight institutional users of version 3 with varying front-end applications to assess factors that support or inhibit migration. Since the release of version 4, Fedora has provided a robust REST-API based primarily on the Linked Data Platform (LDP, https://www.w3.org/TR/ldp/) for resource management operations (Create, Read, Update, and Delete). The services offered and the standards with which they comply are formally documented and versioned separately from the underlying application. This provides stability and predictability while also unlocking the potential to replace the underlying software application with a different implementation that satisfies different use cases. The approach benefits repository migrations because tools can be written against a particular version of the API, which will change less often than the software application. The primary disadvantage to API-based migrations is performance. Transferring large amounts of data over HTTP can be slow. While the API specification does make migrations easier from a tooling perspective, the difficulties of moving large amounts of data will only

---

[3] Examples, all retrieved 2/18/2020: [1] S. Van Tuyl, J. Gum, M. Mellinger, G. L. Ramirez, B. Straley, R. Wick, & H. Zhang. "Are We Still Working on This? A Meta-Retrospective of a Digital Repository Migration in the Form of a Classic Greek Tragedy," *Code4Lib Journal* (issue 41, 8/9/2019), https://journal.code4lib.org/articles/13581. [2] A. Neatrour, J. Myntti, M. Brunsvik, H. Maringanti, B. McBride, & A. Witkowski, "A Clean Sweep: The Tools and Processes of a Successful Metadata Migration," *Journal of Web Librarianship* (11:3-4, 194-208), https://www.tandfonline.com/doi/full/10.1080/19322909.2017.1360167. [3] H. Gilbert & T. Mobley, "Breaking Up With CONTENTdm: Why and How One Institution Took the Leap to Open Source," *Code4Lib Journal* (issue 20, 4/17/2013), https://journal.code4lib.org/articles/8327.
[4] Bridge2Hyku Project Team, *Digital Collections Survey Report* (7/5/2018), p.4, retrieved 2/18/2020 from https://uh-ir.tdl.org/bitstream/handle/10657/3307/Digital%20Collections%20Survey%20Report.pdf?sequence=3&isAllowed=y.

become greater as the amount of data managed by an institution grows.

With Fedora 6, which is the version that will be the focus of the Fedora 3 migrations proposed here, migrations will be easier since the data will only need to be exported, then changed in place without needing to be imported into a new system. This is due to Fedora 6's implementation of the Oxford Common File Layout (OCFL at https://ocfl.io/), "an application-independent approach to the storage of digital information in a structured, transparent, and predictable manner. It is designed to promote long-term object management best practices within digital repositories." Storing Fedora repository resources as OCFL decouples the structure of files from the applications that manage them. Files can be left in place instead of migrated and new software applications, like Fedora 6, can implement the OCFL specification in order to manage the files. The OCFL also addresses Fedora 3 users' concerns around preservation persistence. In the future, newer versions of Fedora will be true software upgrades since they will conform to the OCFL and won't require an export or import at all. Additionally, data stored as OCFL will be easier to migrate to other platforms.

During review of Fedora 3 repositories, variance and customization among front-end applications and data models were found to have the largest impact on the resources required to conduct a migration. Fedora 3 repositories with an Islandora front-end tend to have more commonality and less customization than those using Samvera, which were much more similar to custom Fedora 3 implementations. In some cases, custom frameworks are available on GitHub, but often they are managed in-house, complicating migration and software upgrade efforts, and increasing the burden on local resources. Analysis of data models showed that there are many commonalities among the types of data managed by each institution and that objects tend to be modeled similarly. Application of metadata standards varies but is less of a challenge to migration than data transformations, such as those required to move XML-based metadata into an RDF-based model.

Since the release of Fedora 4, several community-supported migration tools have been developed, but each serves a limited set of use cases. Existing migration tools all require additional work to serve the full range of Fedora 3 implementations, and they focus only on content and metadata migrations. None assist with updates to front-end applications. Two that show promise for the audience of this project are migration-utils (front-end agnostic, most useful for custom Fedora 3 repositories) and migrate_7x_claw (most useful for Islandora repositories). While content and metadata migration tools were identified as being potentially useful, the planning grant focus group participants discussed other barriers that could not be mitigated by tooling, such as transitioning data models. Fedora does not prescribe or enforce data models, so implementers are looking for guidance on how to transition existing Fedora 3 data models to Fedora 4 or higher.

Two themes emerged during the planning project: effort and value. Any software upgrade or migration that requires significant effort needs to provide enough additional value to justify the cost. A software upgrade creates a decision point – to go forward with the new version, stay on the current version, or move to another platform – and the answer depends on value. The new version of the software needs to provide enough additional value compared to both the current version and other software applications to justify the cost of the upgrade and migration. Many survey and focus group respondents noted that decisions to upgrade are made slowly within their institutions. Some said that they are unlikely to migrate off their current system until a critical issue forces them to move. Almost half (48%) of survey respondents have assessed alternatives to Fedora; of those, 20% decided not to move and only 17% elected to move. Overall the survey indicates that three-quarters of the Fedora community does not plan to move to a different repository platform, and most of the rest are undecided. There is value in Fedora 6 that can be realized through an upgrade and migration, provided that tools, documentation, processes, and paths can be produced to make it less costly and more efficient. Este Pope, Head of Digital Collections at Amherst College Library, notes in her letter of support: "We participate in Fedora both because the software helps meet our access and preservation needs for our digital collections content, but also because we get the expanded value of a network of technical expertise and collegial discussion on how best to implement and maintain digital platforms" (Supportingdoc2). To build upon the value of the Fedora network, significant input from the community has been incorporated into design and development of Fedora 6, including findings from the planning project.

Project Design

Project Goals, Assumptions, and Potential Risks

The goal of this project is to enable and provide support for upgrading and migrating from Fedora 3 to Fedora 6. Over 18 months (September 2020 through February 2022), the Fedora Program team at LYRASIS will collaborate with community members to develop, pilot, document, and evaluate tools and paths to improve the migration process for Fedora 3 repositories through two pilots. Improved support for Fedora 3 migrations will increase efficiency for libraries and archives, reduce obstacles, and encourage implementation of an up-to-date system in Fedora 6.

This project assumes that the majority of Fedora 3 repositories would be interested in upgrading to Fedora 6. Data gathered during the planning grant indicates probable interest if the value of Fedora 6 makes it worth the effort and if the resources required for such a migration could be reduced through improved tools, better documentation, training, and defined paths for planning, implementing, and validating. The concerns and needs of institutions still using Fedora 3 have been incorporated into the design and development of Fedora 6. The opportunities Fedora 6 provides around preservation (transparent persistence), querying, fixity, performance, and scale enhance its value, especially in combination with the changes in versions 4 and 5 from which Fedora 3 users have yet to benefit (for example, a standards-based API and native support for linked data). The Fedora Program is reaching out to Fedora 3 users to communicate the importance of upgrading to avoid stability and security issues, and the value of Fedora 6 in providing services not resident in Fedora 3.

The primary risk to the project is if the Fedora 6 public release falls behind schedule. Three development sprints have been completed and an early release is being tested. The team of 4-5 community-based developers anticipates dedicating one week per month to achieve a beta release later in 2020. Feedback from the beta release will be incorporated into the public release, planned for March 2021. The timing of this in relationship to work proposed in this project is noted in the Schedule of Completion. The Fedora community has a well-defined plan for version 6, a highly engaged community of developers and pilot institutions, and a committed Leadership Group that will be engaged in beta testing. For additional details about this, see the design plan at the Fedora Repository Wiki. A delay in the public release may lead to some migration tools and paths beginning development in the beta release.

Project Team and Collaborators

David Wilcox, Fedora Program Leader, will serve as the principal investigator (PI). He has led the program since 2014. He works closely with governance and an international community of stakeholders to secure funding, set priorities, and manage the Fedora program, budget, and in-kind contributions. Wilcox will oversee all aspects of the project, including planning and supporting pilots, documentation of processes and results, communications, testing and evaluation, and outreach and educational activities. His estimated time commitment averages 50% over all 18 months. Wilcox works under the advice and direction of the Fedora Leadership Group and its Steering Group, which will provide oversight and guidance during the project through their regular meetings. To enable him to focus his time on the project, outreach staff from other LYRASIS programs will temporarily take over community management, communications, and membership work with direct support from the Leadership Group.

Andrew Woods will provide overall technical support for project work, especially as related to Fedora 6. He has served as Fedora Technical Lead since 2013 and an editor for the OCFL since 2017. He is responsible for coordinating community efforts around the design, development, and infrastructure of the Fedora platform. During the project, Woods will assist in validation and testing for migrations and upgrades, provide support to the pilots, participate in planning and outreach, and teach at the migration camp. His work over 18 months is estimated at 10%, with most of his time after the Fedora 6 release.

Daniel Bernstein, Fedora Senior Software Engineer at LYRASIS, will serve as the project's Developer. He will provide migration and front-end upgrade support for the custom repository pilot, create a validation tool to compare data between Fedora 3 and 6 repositories to verify that no information is lost, and assist with documentation. Bernstein has provided software development for Fedora since 2011. His time estimate for the project averages 25%.

While Bernstein supports the custom pilot, a subcontractor will supplement Fedora staff capacity and assist with the Islandora pilot site's upgrade and migration, which includes moving from Islandora version 7 to 8, required for moving off the Fedora 3 platform. The subcontractor will assist with data modeling and migration, configuration of Islandora-based migration tools, installation and configuration of the new Islandora 8 repository, and documentation. A "Request for Proposals" was distributed in February 2020 (see Supportingdoc3); preliminary bids have been submitted by two contractors and are used for the budget estimate. A final call for bids will be made if the project is funded, with contracts then established after evaluation of contractors' experience, knowledge of the relevant applications, and cost. Support from the contractor will enable the pilot migrations to move forward at a faster pace.

The project will be implemented with two Fedora 3 institutions representing the majority of front-end application types, Islandora (54% of Fedora 3 users) and custom (38%). The third group, Samvera users, is most like a custom front-end as they use a similar variety of applications and data models to support their repositories; they will be served through tools and paths developed through the custom pilot. The two pilot sites are Whitman College, a liberal arts college, and the University of Virginia, a public research institution. They were selected following interviews to ensure that they are representative of many Fedora 3 repositories in their group; hold diverse types of content in their repositories; and have the capacity, ability, and willingness to upgrade during the project period. Letters of commitment from each are included in Supportingdoc1. Pilot host sites are contributing their staff time to plan, support, and evaluate the upgrade and migration; in return, they will receive consultation and assistance from the Fedora Program team and contractor. This significant contribution of pilot sites' time is not reflected in the budget (to minimize administration cost on their end), but will be tracked to include in documentation.

- Penrose Library at Whitman College uses an Islandora 7 front-end for its Fedora 3 repository. ARMINDA hosts 3 TB of materials including student honors theses, syllabi and course work from Whitman classes, archival photograph and manuscript collections, digitized student newspapers, and oral histories. Penrose Library currently hosts its repository and digital collections on Amazon Web Services (AWS), and the repository is maintained by vendor Born Digital. Internally, the digital collections are supported by a metadata librarian and systems librarian; the Library does not have an in-house developer. Whitman College is part of the Islandora Collaboration Group (ICG), a consortium of colleges and universities that supports members through development, resource pooling, and advocacy. Alzada Tipton, Provost and Dean of the Faculty at Whitman, notes in her letter of commitment: "Our colleagues at ICG schools and in the wider Islandora community who also all face similar migration processes, stand to benefit very materially from our participation, as the documentation we will help to produce and the experience we will be able to share will be instrumental in their ability to also navigate the transition to a new system." Due to the hosted nature of their repository, Whitman College will incur temporary infrastructure costs for additional AWS server space during the migration and upgrade process. This is included in the budget for 12 months.
- The University of Virginia Library uses a custom front-end for a Fedora 3.2.1 repository storing digital collections created circa 2000-2006. Holdings are about 180 GB of Early American Fiction, Modern English, and images of local Art & Architecture. The repository holds high definition images and provenance metadata for preservation. Images are discoverable through the Library's Virgo general portal and web-accessible derivatives are served through the Library's IIIF server and viewer. In addition to supporting preservation, the repository is the source of high definition image orders, of

which the Library receives around a dozen per week. In her letter of commitment, Robin Ruggaber, Director of Strategic Technology Initiatives & Partnerships, notes: "Our experience is that migration tooling and early testing are critical to the success of community wide migration, stabilization of the repository platform, and sustainability of the project." The Library has adequate local server capacity to manage the migration.

The Fedora community will be invited to review and test tools and documentation. While availability of project outputs will be widely communicated within the Fedora community as they become available, the PI also will identify potential test sites during the two pilots and reach out as documentation and tools are produced to encourage adoption. The project team will provide support for those using the project resources to upgrade/migrate their Fedora 3 repositories, including assistance with planning, troubleshooting, and testing. These early adopters will be included in project evaluations.

Project Activities

The project is divided into three phases: Phase 1 involves planning, implementation, tooling, and documentation for the two pilots at Whitman College and the University of Virginia Library. Phase 2 focuses on outreach and support for early adopters of the toolkit, development of educational materials, and continued evaluation and iteration of the toolkit resources. Phase 3 includes training webinars, a hands-on migration camp (workshop), conference and user group presentations to disseminate results, and final project evaluations and reports.

Phase 1 (September 2020 through May 2021, 9 months)
The PI will gather information from pilot sites to inform a kick-off meeting with each to start the project. The first pilot will begin in September 2020, the second about 2-3 months later. Each planning meeting will take place at the pilot location, and include the PI and Tech Lead. The goal of each planning meeting is to define the desired end state for each pilot, map the process and timeline, determine key milestones, verify roles and responsibilities, and define deliverables. For data migration, this includes metadata clean-up and improvements as agreed to and scoped by each pilot's stakeholders. For the Islandora front-end, this includes interface updates and workflow improvements. For the custom front-end, deliverables include updated Fedora 3 to 6 API documentation and support. In general, the desired end states for both pilots are: (1) all data migrated from Fedora 3 to Fedora 6, including mapping of data models and metadata; (2) front-end applications updated to work with the Fedora 6 API; and (3) appropriate performance results for institutions and their repository users. There are differences in approach for the two pilots based on their Islandora or custom front-ends, as noted below. The following steps will occur after each planning meeting.

- The PI will document each pilot project plan and upgrade process. Details and recommendations within the documentation will be updated and expanded as the process unfolds, and will include explanations of the planning and decision-making steps, methods used, lessons learned, and the recommended approach for future adopters.
- For custom Fedora 3 data migration, the project staff will review a representative set of sample data and make a spreadsheet containing all content types and datastreams. Each content type and datastream in Fedora 3 will be mapped to Fedora 6. For the Islandora data migration, all metadata fields in Fedora 3 will also be mapped to fields in Fedora 6. The field mappings will be determined with reference to community best practices and consultation with the pilot. The process of reviewing sample data and community best practices, selecting appropriate mappings, and completing the mapping will be documented in terms of process and guidelines for future use. This

step is often one of the most time-consuming aspects of a migration, so the guidelines will ultimately help other community members make decisions.

- The Tech Lead and Developer will complete the development of a tool that can validate the output of the migration utilities against original Fedora 3 data. This tool will ensure that all of the Fedora 3 content has been successfully migrated to Fedora 6 without data loss. This is an important step to ensure that the migration is a success. The validation tool will be tested through an export of sample content from Fedora 3 using built-in functionality, which will then be configured with the migration tool for source and destination. Results will be reviewed and validated. Tools will be updated as needed until the output is deemed valid, at which time the full migration will be completed and substantiated using the updated migration and validation tools.

- For the custom front-end upgrade, the Developer will document the mapping between Fedora 3 and Fedora 6 APIs. This will involve creating a table that shows how each function of the Fedora 3 API maps to a related function in the Fedora 6 API so that implementers understand how to update their front-end interfaces to work with Fedora 6. Based on the API mappings, front-end applications will be made compatible with Fedora 6 through re-writing API-based interactions; this will be done by the pilot's developers with support from the project team. For Islandora, the site theme changes significantly between Islandora 7 and Islandora 8, so a new one will be created for the pilot by a subcontractor and branded accordingly. The theme controls the look and feel of the site (colors, block placement, etc.). This process will be documented for other Islandora upgrades.

- Following front-end upgrades, the PI will work with local administrators and users (faculty, students, etc.) to test all the expected functionality of the upgraded repository. This includes creating and managing content, search and discovery, and common workflows. Updates to the interface will be made in response to test feedback. Expectations around performance and scale will be documented, and then a test environment will be provisioned to conduct these tests according to the needs of the pilot repository. Results will be captured and reported.

- Following conclusion of testing and any necessary adjustments to align performance with pilot expectations, the PI will work with pilot partners to document their experience in detail to share with the community. This is an important step that will allow similar institutions to better understand the level of effort, timeline, and resources required to do an upgrade to Fedora 6.

- Drawing from all documentation and tools produced through both pilots, the PI will create a toolkit that community members can use to conduct their own migrations and upgrades to Fedora 6. It is expected that some tools and documentation would apply to all Fedora 3 users independent of their front-end application, and others would be specific to a front-end application. Tools and documentation for both pilots will be provided under open source or Creative Commons attribution licenses as relevant. The toolkit will be shared through the Fedora GitHub and Confluence wiki spaces, and sustained and supported as part of the Fedora community.

Phase 2 (June through September 2021, 4 months)

Following its release in spring 2021, the PI will lead outreach to promote availability of the Fedora 3 migration/upgrade toolkit. Institutions with known Fedora 3 installations will be contacted and asked to try out the toolkit and provide feedback. This outreach will be based on the Fedora registry and participation in the survey during the planning grant. The project team will be available to respond to questions, provide support, and assist with troubleshooting for these early adopters of the toolkit. This may include scheduling calls with implementers, hosting open office hours, and providing a dedicated Slack channel for communications. Feedback from community members will be used to update the toolkit. Training materials also will be developed during phase 2 for the migration camp and webinars scheduled to occur in Phase 3.

Phase 3 (October 2021 through February 2022, 6 months)

A dedicated Fedora 3-to-6 migration camp will be held after the toolkit is released, in fall 2021. The goal of the 2.5-day camp is to allow participants to work through a test migration using the toolkit, with their own test data if provided. A draft curriculum is included in Supportingdoc4. The audience for the camp is developers, data migration specialists, and technology leads responsible for managing and supporting their library's or archives' Fedora 3 installations. Registration will be free, and travel funding will be provided through the grant for one person each from up to 20 Fedora 3 institutions (additional participants from a single institution would be allowed, without grant travel support, for up to a maximum of 30 in the camp). Based on past experience with Fedora Camps, four instructional leaders will be involved to work directly with teams of participants, make presentations, share experiences, and assist with questions and troubleshooting. The camp will be hosted at a venue provided for free by a Fedora institutional user. Outreach during phase 2 will promote the camp and grant support for travel costs will encourage participation.

Based on the training materials developed for the in-person camp, two shorter webinars will be developed and offered by the PI to share recommendations and results from the toolkit more broadly. One webinar will focus on an Islandora upgrade and migration, the other on a custom repository. Each webinar will be offered during the final phase of the grant period, and recordings with slides from each session will be posted with the toolkit for those unable to attend. The webinars will be especially useful to those considering or just beginning to plan an upgrade of their Fedora 3 repository, as they will walk through the process and connect attendees to the toolkit resources that would be most useful at any stage.

To share the results of the project more broadly, both within and outside of the Fedora community, the PI will make presentations at two [Fedora User Group meetings](#) and three national conferences, such as those sponsored by Code4Lib, the Coalition for Networked Information, Open Repositories, the Digital Library Federation, and Islandora. These presentations may incorporate co-speakers from pilot or early adopter sites. Presentations within the Fedora User Groups will focus on the toolkit and the technical outcomes from the project. Presentations through national conferences will speak more broadly about what was learned, through both the planning grant and this project, in regard to requirements and impact of major software upgrades, especially those requiring data migration.

Evaluations

During the pilots, feedback will be gathered from the host stakeholders on an ongoing basis to compare outputs with the plans developed at the beginning of the projects; this will primarily occur through regular phone/web meetings of stakeholders throughout the process. Feedback will be incorporated into recommendations included in the toolkit. In addition, as noted above, validation tools will be developed to ensure effective and accurate data migration, and testing will be conducted on both pilot repositories to ensure performance and scale meet stakeholder expectations.

Participants in the migration workshop, toolkit webinars, and Fedora User Group events will be asked via survey for feedback on each educational program. This will include gathering information about plans and/or status of Fedora 3 migrations/upgrades to Fedora 6. A final survey will occur toward the end of the project focused on known Fedora 3 institutions at the beginning of the grant (fall 2020). This will gather information and feedback about the value of the toolkit and community-wide status of upgrades/migrations. Project indicators of success are:

- Pilot Fedora 6 repositories are functioning to meet stakeholder expectations, and pilot staff involved in the upgrade and migration assess themselves as able to manage the new repository environment.
- A majority of U.S. Fedora 3 institutions have accessed the toolkit as reported through the final project survey.
- A decrease in the percentage of the Fedora 3 community that is undecided about migrating (63%

from the planning grant survey) and an increase in the percentage that plan to upgrade and migrate to Fedora 6 (using the baseline of 20% from the planning grant survey). This would be measured in the final project survey.

Sustainability

Documentation and tools produced through this project will be maintained, updated, and sustained by the Fedora community through its governance, development, communications, and planning infrastructure. Fedora's international [membership](#) actively supports the software with leadership as well as financial and development resources. With the availability of the migration toolkit for Fedora 3 users, the program staff and Fedora Leadership will continue to encourage repositories to upgrade to supported versions of the application post-grant, for as long as needed. Efforts will include continuing webinars and training sessions, providing advice and support for repositories during the upgrade/migration process, and ongoing additions and updates to documentation, case studies, and migration tools.

National Impact

Continued reliance on an unsupported software version, like Fedora 3, presents many risks to an organization. It limits interoperability with newer technologies that can add or improve functionality for users. It eliminates the advantages of community support – having a team of experts to contribute to updates and address issues – which is now focused on Fedora 4, 5, and 6. Stability and security decrease as maintaining and updating Fedora 3 falls on an institution's own staff. Nearly 60% of planning grant survey respondents had 3 or fewer staff supporting their repository. Maintaining an old software application so that it works securely on new systems will take increasing amounts of time as the rest of the technology supporting a library or archives continues to advance. Repository performance decreases, and scale becomes an issue as digital collections continue to grow. In the end, access to valuable digital collections may be limited or lost, impacting teaching, learning, and research.

Limited staff resources are a key obstacle to upgrading. The pilot project proposed here seeks to lessen the demand on staff resources (although it will not eliminate it) by providing tools, documentation, best practices, case studies, and training materials to assist a library or archives in planning for and implementing an upgrade. Continued focus from the Fedora Leadership and Steering Groups as well as program staff on the value of Fedora 6 will further encourage upgrades by addressing feature and functionality issues that may have kept some Fedora 3 users from upgrading. An upgrade and migration decision for a Fedora 3 repository becomes easier to justify when the value of Fedora 6 is communicated in combination with the availability of a toolkit that can help reduce the level of effort. As more Fedora 3 repositories upgrade and migrate, additions to the toolkit (case studies, documentation, data model maps, etc.) will expand resources to assist others.

Planning for and implementing a repository upgrade and migration takes time, especially with limited resources to dedicate to the effort. While pilot institutions will each complete a repository upgrade and migration during the project period (18 months), early adopters of the toolkit may still be planning or in early stages at the end of the project. Long-term, all Fedora 3 repositories will eventually upgrade or be replaced by new platforms. While the toolkit will be designed specifically for moving from Fedora 3 to 6, it can nonetheless provide a useful model for Fedora 3 institutions moving to different platforms. It also can serve as a model for supporting future Fedora upgrades, providing processes, workflows, data maps, and documentation structures that can be adapted for new situations. The key theme of effort and value defined during the planning grant will influence work during the pilot project and, long-term, the future evolution of Fedora. While this project focuses on decreasing the effort of the Fedora 3 upgrade and migration, it also provides value as a model to others. Ultimately, those who use and depend upon repositories to preserve and openly share scholarly publications, research data, and cultural heritage will benefit through improved and expanded access to content in upgraded, up-to-date Fedora repository platforms.

Fedora Migration Paths and Tools: A Pilot Project
Schedule of Completion

Task	Assigned To	9/20	10/20	11/20	12/20	1/21	2/21	3/21	4/21	5/21	6/21	7/21	8/21	9/21	10/21	11/21	12/21	1/22	2/22
Phase 1																			
Public release of Fedora 6	Provided for reference							▮											
Gather pilot information	PI, Pilots	▮																	
Meet with each pilot team	PI, Tech Lead, Pilots	▮	▮																
Document project plan of each pilot	PI, Pilots	▮	▮																
Document the upgrade process	PI, Developer			▮	▮	▮	▮	▮											
Review sample data (UVa)	PI		▮																
Map data models and metadata (UVa)	PI, Pilot		▮	▮															
Document the process (UVa)	PI		▮	▮															
Apply methodology to second pilot																			
Review sample data (Whitman)	PI, Contractor				▮														
Map data models and metadata (Whitman)	Contractor, PI, Pilot				▮														
Document the process (Whitman)	Contractor, PI				▮	▮													
Complete validation tool	Tech Lead, Developer				▮	▮													
Run test migration (UVa)	Developer					▮													
Update migration tools (UVa)	Developer					▮													
Complete data migration (UVa)	Developer						▮												
Document the process (UVa)	PI, Developer					▮	▮												
Apply methodology to second pilot																			
Install Islandora 8 (Whitman)	Contractor					▮													
Configure migration tool (Whitman)	Contractor					▮													
Run test migration (Whitman)	Contractor					▮													
Update migration tool configuration (Whitman)	Contractor					▮													
Complete data migration (Whitman)	Contractor						▮												
Document the process (Whitman)	PI, Contractor					▮	▮												
Submit conference proposals to relevant events	PI						▮	▮	▮	▮	▮								
Document Fedora 3 vs Fedora 6 API (UVa)	Developer						▮												
Support front-end application updates (UVa)	Developer. Pilot						▮	▮											
Create a new Islandora 8 theme (Whitman)	Contractor							▮											
Configure the repository (Whitman)	Contractor							▮											
Test and update the repository (UVa)	PI, Developer							▮	▮										
Test performance and scale (UVa)	Developer							▮	▮										
Apply methodology to second pilot																			
Test and update the repository (Whitman)	Tech Lead, Contractor								▮	▮									
Test performance and scale (Whitman)	Tech Lead, Contractor								▮	▮									
Produce case studies	PI									▮	▮								
Schedule webinar featuring pilots	PI									▮									
Create toolkit	PI, Developer									▮									
Phase 2																			
Share draft toolkit with community for feedback & evaluation	PI										▮								
Perform targeted outreach to specific institutions	PI										▮								
Respond to feedback and iterate on toolkit	PI, Tech Lead											▮	▮						
Schedule a migration workshop/camp event	PI											▮							
Develop training materials for migration workshop/camp	PI, Tech Lead											▮	▮						
Develop online training resources	PI, Tech Lead												▮	▮					
Develop conference presentation slides	PI												▮	▮					
Phase 3																			
Revisit pilots	PI														▮				
Report on work at conferences	PI														▮	▮			
Present on toolkit at user group meetings	PI														▮	▮	▮		
Host migration workshop/camp	PI, Tech Lead															▮	▮		
Deliver online training webinar	PI, Tech Lead																▮	▮	
Draft final report	PI																	▮	▮

Pink = Pilot 1 at University of Virginia
Blue = Pilot 2 at Whitman College

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.

The digital content, resources, and assets created as deliverables and work products will be under the copyright of LYRASIS and released to the public using a CC BY 4.0 license . This license is applied to all Fedora wiki content and documentation. There are no conditions on access or use.

Software created as part of the project (potentially for migration tools, validation tool, and repository front-ends) will be under the copyright of LYRASIS and released to the public using the relevant software program's open source license: for Fedora, Apache 2.0; for Islandora, GNU General Public License; for migration-utils, Apache 2.0; for migrate-7x-claw, MIT 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.

LYRASIS will assert no ownership rights over new digital products produced during the project other than those defined by the relevant licenses. There will be no conditions on access or use.

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.

While we do not anticipate producing digital products that would involve privacy concerns during the project, we will gather express, written consent in advance from the pilot organizations to document their migration and upgrade processes. Evaluation results will be made anonymous prior to inclusion in any reports that are to be publicly released.

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.

Our project will produce documentation for each of 2 pilots, including processes, data models and maps (by content type and metadata fields), API maps between Fedora 3 and Fedora 6, and case studies of plan implementations and outcomes. Training materials, including curriculum, exercises, and slides, will be produced for 1 workshop/camp , 2 Fedora user group meetings, 1 webinar (potentially held multiple times), and 3 conference presentations. Materials will be produced in Google formats for easy sharing and editing (Docs, Sheets, Slides). Final publications may exist in these formats and/or as PDFs.

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.

Digital content will be produced in Google Docs, Sheets, Slides and/or Forms.

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.

We plan to use the following formats: Google Forms, Docs, Sheets, Slides, as well as .PDF. We can provide other formats on request, such as .txt, .docx, or pptx.

Workflow and Asset Maintenance/Preservation

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

Google Docs and Sheets tracks versions of the assets, including user and time stamps of edits that will assist the P.I. and project team in oversight of quality control and workflow for digital content. Pilot participants will be able to edit final versions of their documentation prior to publication. The P.I. will engage other LYRASIS staff in proofreading for drafts of content to be published.

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).

Final copies of digital content/assets produced through the project will be uploaded to the Fedora wiki and websites as well as to LYRASIS' DuraSpace DuraCloud account for duplicated preservation storage. LYRASIS commits organizational funding for maintaining these resources, which are central to operations for the Fedora community. Each pilot's documentation will be deposited into their own institution's repository for discovery.
Technical documentation for migration tools will be maintained in the relevant GitHub repository. Technical documentation for the validation tool will be maintained in the Fedora GitHub repository.
All content will be stored using open standards to ensure long-term access.

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).

Content stored on the Fedora wiki and GitHub will be well-described to facilitate indexing in popular search engines (e.g. Google, Bing) which is one of the primary ways the content will be discovered. Technical metadata will be automatically generated when content is uploaded to DuraCloud. Pilot documentation deposited into local Fedora repositories will have automatically generated technical and administrative metadata.

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

Metadata and content uploaded to DuraCloud is automatically duplicated and integrity checked bi-annually.

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).

We will present the findings on this project through conferences, Fedora user group meetings, a Fedora migration workshop/camp, and a webinar. In those presentations we will point to the locations of the digital assets by weblink. In addition, the DuraSpace/Fedora wiki and website is crawled by search engines so assets shared there will be discoverable via Google, Bing, etc.

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).

The digital content will be made openly available through the Fedora website and wiki. It will be accessible via standard web browsers and provided in non-proprietary formats (PDF, Google Docs/Sheets/Slides) that do not require fee-based software to access.

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.

Designing a Migration Path for Fedora project documentation and report at
https://wiki.lyrasis.org/display/FF/Designing+a+Migration+Path.

The Fedora Digital Preservation Survey report at
https://duraspace.org/fedora/resources/publications/fedora-digital-preservation-survey/.

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.

> migration-utils: this is a configurable command-line tool that transforms Fedora 3 data into Fedora 6 data. It will primarily be used by institutions with custom Fedora 3 repositories when they migrate to Fedora 6.
>
> migrate_7x_claw: this is a configurable Drupal-based tool that converts Islandora 7 data to Islandora 8 data. It will primarily be used by institutions when they migrate from Islandora 7 to Islandora 8.
>
> Validation tool: this is a command-line tool that compares Fedora 3 data with Fedora 6 data to ensure nothing was lost or corrupted after a migration. The primary audience will be institutions that have recently completed a Fedora 3 to 6 migration.

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.

> Both existing migration tools will be used and updated as needed during this project. They are fit for purpose but have not been used widely and will likely need to be modified as specific use cases are encountered.
>
> We are not aware of any existing tools capable of comparing and validating Fedora 3 and Fedora 6 data following a migration. The validation tool we will create will be specific to this use case.

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.

> migration-utils and the validation tool are written in Java. migrate_7x_claw is written in PHP. These languages correspond with the languages Fedora and Islandora are written in. Developers will develop code locally using their preferred IDEs, and GitHub will be used to facilitate collaboration, version control, and the long-term preservation of the code. GitHub is the global standard for collaborative code management.

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

migration-utils will bridge the gap between Fedora 3 and Fedora 6 by converting the data into a format compatible with the new version of the Fedora software. The validation tool will build on this work by providing a means of verifying that a migration has been completed successfully with no data loss or corruption. migrate_7x_claw is built on the Drupal Migrate API and extends the standard Drupal 7 to Drupal 8 migration functionality to support the additional requirements of a full Islandora 7 to Islandora 8 migration.

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

migration-utils runs on the command line and requires Java 11. The validation tool will also run on the command line and require Java 11. migrate_7x_claw is a Drupal 8 module that must be installed in Islandora 8 and requries a running Islandora 7 instance as its source. It also requires the Drush module for Drupal 8.

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

The Feodra Guide for New Developers describes the development process we will follow for migration-utils and the validation tool: https://wiki.lyrasis.org/display/FF/Guide+for+New+Developers. For migrate_7x_claw we will follow the (similar) Islandora guidelines for contributing code: https://islandora.github.io/documentation/contributing/CONTRIBUTING/. We will create documentation as we develop the code and this documentation will be maintained and updated over time to coincide with any changes in functionality or requests from users. Anyone can contribute to the documentation either on the Fedora wiki or on GitHub.

B.5 Provide the name(s), URL(s), and/or code repository locations for examples of any previous software your organization has created.

Fedora Repository Home (wiki) at https://wiki.lyrasis.org/display/FF/Fedora+Repository+Home

Fedora 5.1.0 https://duraspace.org/fedora/download/

Code repositories for Fedora and related modules: https://github.com/fcrepo4, https://github.com/fcrepo4-exts

Access and Use

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

> The software will be made available through the Fedora wiki and announced on community mailing lists. The source code will be made avaialble on GitHub.

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

Name of publicly accessible source code repository:

> fcrepo4-exts (migration-utils and validation tool); Islandora-Devops (migrate_7x_claw)

URL:

> https://github.com/fcrepo4-exts; https://github.com/Islandora-Devops

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.

> Not applicable to this project.

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?

Not applicable to this project.

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.

Not applicable to this project.

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

Not applicable to this project.

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?

Not applicable to this project.

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

Not applicable to this project.

A.7 Identify where you will deposit the data:

Name of repository:

Not applicable to this project.

URL:

Not applicable to this project.

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

Not applicable to this project.