

The University of Oregon (UO) Libraries, in partnership with Oregon State University Libraries and Press (OSULP), seeks \$203,508 from an IMLS National Leadership Grant to support its project, **Open Impact: Developing Robust Analytics for Open Source Solution Bundle Hyrax**. This one-year project grant addresses the development of a suite of analytics and reporting tools for Hyrax, an open source, Samvera-powered repository front end.

Focusing on creating new dashboard functionality to enable repository managers to tell the story of their collections and their users, this project will help reduce dependency on third-party, for-profit repository platforms and demonstrate the value of repository services to our stakeholder communities. This work will contribute directly to Hyrax's community development initiatives, with all code contributions, documentation, and project updates made openly available and coordinated with platform stakeholders across the country in alignment with identified Samvera community needs. With Hyrax forming the code base of Hyku, the turnkey repository platform ([LG-70-15-0006-15](#), [LG-70-17-0217-17](#)), this project will address the need for analytics, identified in the HykuDirect Pilot Gap Assessment as "critical to be competitive in the repository space" and a "must have" for future development.

With library backlash against the acquisition of bepress by Elsevier in August 2017 and the IMLS-funded Bridge2Hyku toolkit ([LG-70-17-0217-17](#)) supporting migration from CONTENTdm to Hyku, repository managers around the country have sought alternatives for their feature-rich platforms. Hyku was identified as a potential turnkey alternative, but the Hyku Pilot Gap Assessment identified a significant number of software gaps that would need to be addressed to permit production-level service. Some of these features have been incorporated into the Hyrax roadmap, and the Hyrax Analytics Working Group, with representatives from thirteen Samvera Partners, has assembled user stories, use cases, and requirements. With a long list of development priorities and limited resourcing, analytics has not yet been the focus of a major development effort and significant time investment would be required to bring Hyrax (and Hyku) up to feature-parity for libraries across sectors, much less improving on third-party, for-profit competitors.

IMLS and the repository community have invested in building best practice and solutions for repository analytics, including the Digital Library Federation's white papers, Institutional Repository Usage Statistics UK (IRUS-UK), and the Repository Analytics & Metrics Portal (RAMP) web service ([LG-06-14-0090-14](#)). Libraries have also begun looking at open source web analytics application Matomo (formerly Piwik) to reduce dependency on third-party, for-profit analytics platforms and to respect user privacy. This project will incorporate existing and emerging best practices for repository analytics with an eye towards balancing accuracy, data privacy, and functionality and contribute to a growing area of library research. Analytics facilitate a variety of reporting functionality for platform administrators, as well as end-users, to understand who and how these platforms and the collections they contain are being used. From an undergraduate student interested in the download count of their honors' thesis to a public library demonstrating the international reach of their special collections, the metrics provided by repository analytics can tell a variety of stories.

These stories and this project directly support the IMLS agency-level goal of *Build Capacity* by enabling repository managers, as well as library administrators, to justify initial and continued investment in digital platforms that ultimately increase public access to research, data, and rare and special collections. This project also aligns with the category of *National Digital Infrastructures and Initiatives* focal area of *Expanding Digital Cultural Heritage Capacity* through continued development of shared tools, as well as addressing barriers to the adoption of those tools, as analytics were designated as a requirement. It also contributes to the *Opening Scholarly Communications* focal area by creating interoperable repository functionality that will support green open access, new models for return on impact, and benchmarking initiatives.

The University of Oregon (UO) Libraries, in partnership with Oregon State University Libraries and Press (OSULP), seeks \$203,508 from an IMLS National Leadership Grant to support its project, **Open Impact: Developing Robust Analytics for Open Source Solution Bundle Hyrax**. This one-year project grant addresses the development of a suite of analytics and reporting tools for Hyrax, an open source, Samvera-powered repository front end. Focusing on creating new dashboard functionality to enable repository managers to tell the story of their collections and their users, this project will help reduce dependency on third-party, for-profit repository platforms and demonstrate the value of repository services to our stakeholder communities. This work will contribute directly to Hyrax's community development initiatives, with all code contributions, documentation, and project updates made openly available and coordinated with platform stakeholders across the country in alignment with identified Samvera community needs. With Hyrax forming the code base of Hyku, the turnkey repository platform ([LG-70-15-0006-15](#), [LG-70-17-0217-17](#)), this project will address the need for analytics, identified in the HykuDirect Pilot Gap Assessment as "critical to be competitive in the repository space" and a "must have" for future development.¹

Statement of National Need

With library backlash² against the acquisition of bepress by Elsevier in August 2017 and the IMLS-funded Bridge2Hyku toolkit ([LG-70-17-0217-17](#)) supporting migration from CONTENTdm to Hyku, repository managers around the country have sought alternatives for their turnkey, feature-rich platforms.³ Hyku ([LG-70-15-0006-15](#)) was identified as a potential turnkey alternative, but the Hyku Pilot Gap Assessment identified a significant number of software gaps that would need to be addressed to permit production-level service.⁴ Some of these features have been incorporated into the Hyrax roadmap.⁵ The Hyrax Analytics Working Group, with representatives from thirteen Samvera Partners, began the work of assembling user stories, use cases, and requirements in summer 2017.⁶ With a long list of development priorities and limited resourcing, analytics has not yet been the focus of a major development effort and significant time investment would be required to bring Hyrax (and Hyku) up to feature-parity for libraries across sectors, much less improving on third-party, for-profit competitors.

Building an analytics module for Hyrax, an open source solution bundle that sits on top of the Samvera repository framework, will impact previous IMLS-funded projects Hyku ([LG-70-15-0006-15](#)) and Avalon ([LG-70-17-0042-17](#)), both of which use (or will use⁷) Hyrax as their underlying codebase. As depicted in the following diagram, the module would likely pull data from third-party analytics platforms and the Solr API:

¹ HykuDirect Pilot Gap Assessment: <https://docs.google.com/document/d/1mpYLS2pqwKDCCxF0Kb-JTC4OjPLgmbJ6d1IBunxHqAw/edit#heading=h.up2kw3izu6dy>

² Inside Higher Ed Announcement: <https://www.insidehighered.com/news/2017/08/03/elsevier-makes-move-institutional-repositories-acquisition-bepress>

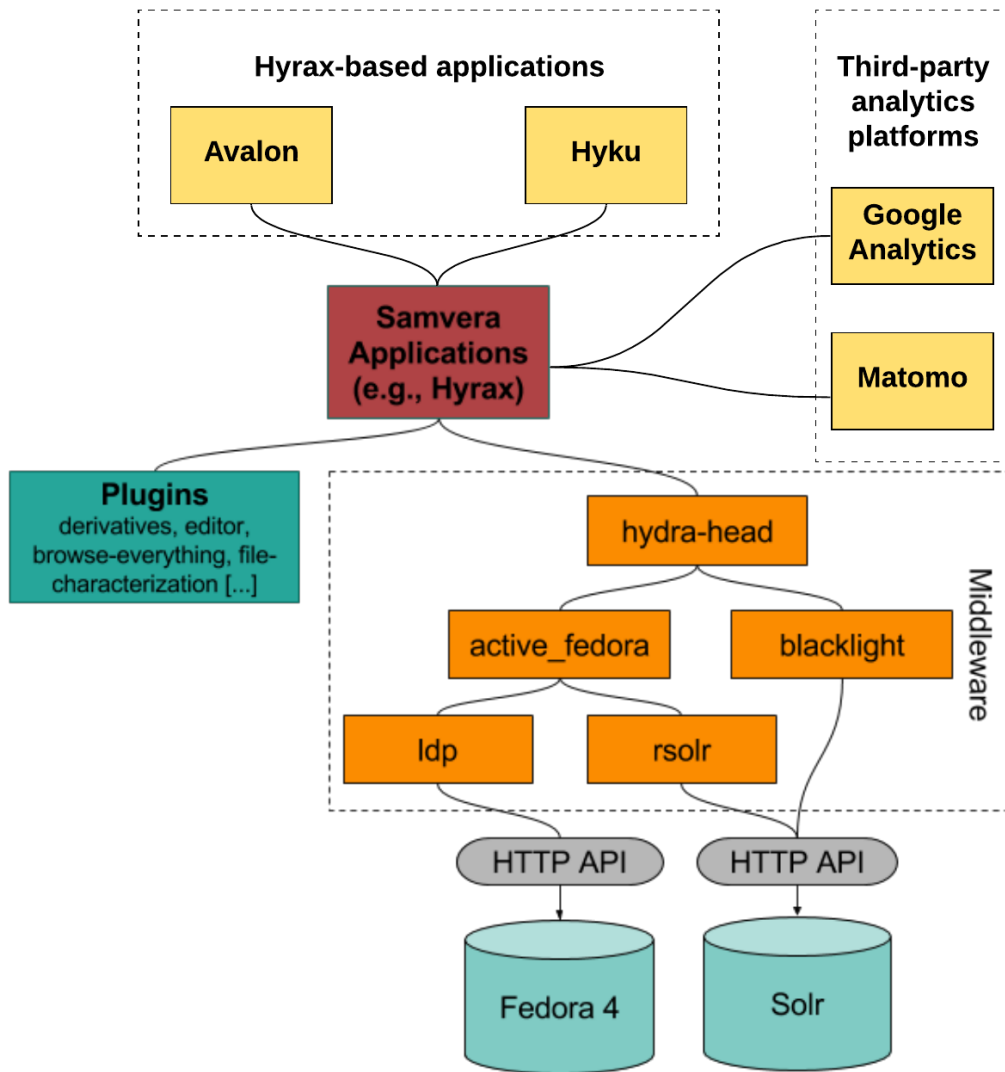
³ BePrexit from University of Pennsylvania: <https://beprexit.wordpress.com/2017/10/13/48-hours/>

⁴ HykuDirect Pilot Gap Assessment, *ibid*.

⁵ Hyrax Roadmap: <https://wiki.duraspace.org/display/samvera/Hyrax+Roadmap>

⁶ Hyrax Analytics Working Group: <https://wiki.duraspace.org/pages/viewpage.action?pageId=87461330>

⁷ Avalon Release Road Map: <https://wiki.dlib.indiana.edu/display/VarVideo/Avalon+Release+Road+Map>



Adapted from a CC BY 4.0 International image by Samvera

IMLS and the repository community have invested in building best practices and solutions for repository analytics, including the Digital Library Federation’s white papers^{8,9}, Institutional Repository Usage Statistics UK (IRUS-UK)¹⁰, and the Repository Analytics & Metrics Portal (RAMP) web service ([LG-06-14-0090-14](#)). Libraries have also begun looking at open source web analytics application Matomo (formerly Piwik) in an effort to reduce dependency on third-party, for-profit analytics platforms and to respect user privacy.^{11,12} This project will incorporate existing and emerging best practices for repository analytics with an eye towards balancing accuracy, data privacy, and functionality and contribute to a growing area of library research. Analytics facilitate a variety of reporting functionality for platform administrators, as well as end-users, to

⁸ DLF Best Practices for Google Analytics in Digital Libraries: <https://osf.io/th8av/>

⁹ DLF Surveying the Landscape: Use and Usability Assessment of Digital Libraries <https://osf.io/uc8b3/>

¹⁰ Institutional Repository Usage Statistics UK: <http://irus.mimas.ac.uk/>

¹¹ Using Piwik Instead of Google Analytics at the Cornell University Library: <https://doi.org/10.1080/0361526X.2016.1245645>

¹² Piwik: Breaking Away from Google Analytics: <http://open-shelf.ca/160215-piwik/>

understand who and how these platforms and the collections they contain are being used. From an undergraduate student interested in the download count of their honors' thesis to a public library demonstrating the international reach of their special collections, the metrics provided by repository analytics can tell a variety of stories.

These stories and this project directly support the IMLS agency-level goal of *Build Capacity* by enabling repository managers, as well as library administrators, to justify initial and continued investment in digital platforms that ultimately increase public access to research, data, and rare and special collections. This project also aligns with the category of *National Digital Infrastructures and Initiatives* focal area of *Expanding Digital Cultural Heritage Capacity* through continued development of shared tools, as well as addressing barriers to the adoption of those tools, as analytics were designated as a requirement. It also contributes to the *Opening Scholarly Communications* focal area by creating interoperable repository functionality that will support green open access, new models for return on impact, and benchmarking initiatives.

Project Design

The University of Oregon will act as fiscal agent for this project with 2 co-PIs: Franny Gaede (UO) and Margaret Mellinger (OSULP). UO and OSULP are long-term collaborators on the [Oregon Digital](#) initiative, which has led Hydra and Samvera development to create a shared repository platform, uniquely positioning us to work with the Samvera community and the contractor hired to complete the analytics modules. The project manager will be Ray Henry (UO); Steve Van Tuyl (OSULP) will coordinate the contractor, Samvera community, and UO/OSULP teams as project manager and current Hyrax product owner; and UO/OSULP teams will contribute user testing (Mary Galvin, UO), front-end design and development (David McCallum, UO; Kenna Warsinske, OSULP), and technical consulting with the contractor (Jeremy Echols and Linda Sato, UO; Josh Gum and Ryan Wick, OSULP).

The project has a proposed timeline of one year, running from September 1, 2019 to August 31, 2020. Work will fall into four phases:

Phase One Activities (September 1, 2019 – November 30, 2019)

- a. Work with full Oregon Digital team to review and update the needs assessment created by the Hyrax Analytics Working Group in 2017:

User stories

Repository Administrators

Repository administrators can report out the number and total size of collections in their repository, including the number of items, files, and file size of those collections. They can report out the total number of unique objects in the repository, the total number of times those unique files were accessed, the total number of queries conducted during the reporting period, and the number of unique downloads for objects in any given collection. Administrators can view the number of times collections were accessed by month, view the number of unique views of an object, and aggregate statistics based on file type or resource type.

Repository administrators can integrate statistics from a variety of sources into Hyrax beyond Google Analytics and Access analytics via a REST endpoint and provide JSON output of most viewed author, most downloaded, and other such information to contribute to a larger scale reporting system such as the Association of Resource Libraries. They can use a calendar widget to generate statistics by month, a group of months, by fiscal year, or any other period, as needed. They can view graphics of item visibility and see items in workflow prior to publication on a dashboard.

End-Users and Depositors

End users can search, browse, or navigate to a collection and, based on the results, view a table and/or figure of page views and downloads for those items and download a file with that data. They can see the Top 100 most-viewed works a thumbnails and number of views. If they choose to create a user account, they can save a search or report as part of their user profile.

Depositors have the same access as end-users, as well as access to statistics about the discovery and use of their works. These statistics include the total number of times the page with their work has been visited, that number over time, the number of unique visitors, a list of referring URLs, geo-location of visitors based on IP or partial IP addresses, the total number of downloads of a file in a work, and a time series of downloads.

Use cases

Administrative Dashboard

On the administrative dashboard, repository administrators will see, at a glance, graphics for the number of collections, including published and unpublished collections, as well as the number of collections over time. This dashboard will show the number of works by work type, resource type, and visibility type; the number of files; the number of new items in the repository over time; the amount of user activity over time; and the number of visitors over time.

Repository administrators will be able to see the number of collections and information about each collection, including the number of works and files in each collection and the number of users interacting with the collection. They will see graphics for the total number of queries run on the repository over time and graphics of visibility that show items and collections in various stages of workflow. All charts on the administrative dashboard will allow for setting a new date range with a calendar widget.

Role-Based Reporting Dashboards

Administrator dashboards will allow administrators to report out the number of downloads based on filetype or resource type, the number of times collections were accessed in a given period, the number of unique downloads for objects in a given collection, and the total number of queries conducted for collections. Administrators will be able to use a calendar widget to select the period for which they would like to generate reporting.

Depositor dashboards will allow depositors to report out on the total number of times the page with their work has been visited, that number over time, the number of unique visitors, a list of referring URLs, geo-location of visitors based on IP or partial IP addresses, the total number of downloads of a file in a work, and a time series of downloads.

Item-Level Statistics

A table and/or chart of pageviews and downloads for the item is available and the data can be downloaded. Other statistical information and alternative metrics, such as citations, Tweets, Mendeley saves, as available, can be viewed as available. The location of downloads can be viewed on a map and a table of places or countries.

Analytics Management

Analytics can be accessed via a REST endpoint and provide JSON output of most viewed author, most downloaded, and other such information to contribute to a larger scale reporting system such as the Association of Resource Libraries. Legacy statistics from previous repository platforms can be integrated and viewed in current dashboards and reports.

Setup and Configuration of Analytics

Repository administrators can easily add their Matomo or Google Analytics account and control its tracking scope in the dashboard.

Analytics Documentation and Best Practices

Integrated into the interface is documentation to guide repository administrators and depositors to help them understand how analytics work in Hyrax and what best practices the Samvera community has to offer on analytics platforms and sources to use and caveats.

Analytics Reports Delivered by Email

Users receive email with consolidated analytics about their deposits (i.e., "Your article was accessed 21 times this month, your data was downloaded 5 times this week"). Depositors receive an email when a particular milestone has been reached (i.e., "Congratulations! your dissertation has been viewed 100 times!"). Depositors have the ability to configure their analytics notifications in their user profile and can turn on or off analytics reports and change the frequency of notification to weekly, monthly, quarterly, etc.

Requirements Tables

- The full project team will evaluate the Requirements Tables to see which Use Cases in the initial minimum viable product, increased functionality, and enhancements sections will make up a reasonable build in Phase 3.
 - Project team will review the [definitions](#) based on Portland Common Data Model (PCDM) and COUNTER for the following terms: Item, Work, Work Type, Resource Type, FileSet, File, File, View, Pageview, Download (for Work and File), Administrator, Collection, Aggregation, Depositor, Investigation, Request.
- b. Primary Investigators Franny Gaede and Margaret Mellinger will complete the University of Oregon institutional Request for Proposal process and hire the contractor who will be responsible for developing the analytics modules. A draft of this RFP is available in our Supporting Documents.
 - c. Franny Gaede, Margaret Mellinger, Steve Van Tuyl, and other members of the project team as available will engage in community outreach online (Samvera Partner Calls, Samvera Virtual Connect).

Phase 2 Activities (December 1, 2019 – February 29, 2020)

- a. The project team will work with the contractor to refine analytics design based on initial mockups and data from selected third-party analytics application(s) (i.e., Matomo, Google Analytics), with consideration for accessibility, accuracy, data privacy, and functionality of the platforms.
- b. Steve Van Tuyl, Ray Henry, and contractor will coordinate requirements with the Samvera community through the Hyrax Working Group, ensuring that any changes made to the Hyrax Analytics Working Group Requirements Tables continue to represent the needs of the community.
- c. Franny Gaede, Margaret Mellinger, Steve Van Tuyl, and other members of the project team as available will engage in community outreach online (Samvera Partner Calls, Samvera Virtual Connect)

- d. A member of the project team will engage with the community and apply to present findings so far at Code4Lib 2020

Phase 3 Activities (March 1, 2020 – May 31, 2020)

- a. The contractor, with input from the project team and community, will develop the analytics modules. Code review will be conducted by Jeremy Echols, Josh Gum, Linda Sato, and Ryan Wick to ensure that the code meets Samvera/Hyrax conventions.
- b. The contractor, with input from David McCallum and Kenna Warsinske, will implement front-end visualization components.
- c. The full project team, led by Mary Galvin, David McCallum, and Kenna Warsinske, will engage in usability testing.
- d. Franny Gaede, Margaret Mellinger, Steve Van Tuyl, and other members of the project team as available will engage in community outreach online (Samvera Partner Calls, Samvera Virtual Connect)

Phase 4 Activities (June 1, 2020 – August 31, 2020)

- a. The contractor, with input from the project team and community, will complete the analytics modules. Code review will be conducted by Jeremy Echols, Josh Gum, Linda Sato, and Ryan Wick to ensure that the code meets Samvera/Hyrax conventions.
- b. The contractor, with input from David McCallum and Kenna Warsinske, will complete implementation of front-end visualization components.
- c. The full project team, led by Mary Galvin, David McCallum, and Kenna Warsinske, will complete usability testing.
- d. Franny Gaede, Margaret Mellinger, Steve Van Tuyl, the contractor, and other members of the project team as available will engage in community outreach online (Samvera Partner Calls, Samvera Virtual Connect)
- e. A member of the project team will engage with the community and apply to present findings so far at Open Repositories 2020
- f. Franny Gaede and Margaret Mellinger will prepare a Samvera community-oriented project report evaluating the success of the project by comparing the initial needs assessment created by the Hyrax Analytics Working Group in 2017 to the feature set ultimately available in the analytics module we create and by the level of adoption and feedback from the community.

Potential Risks

There are a few potential risks for this project. Timeline dilation around institutional processes for the Request for Proposal for the contractor is a concern, impacting who may be available to hire and the timeline for deliverables. However, UO has recently navigated these processes and have staff at both universities, including Franny Gaede, Margaret Mellinger, Ray Henry, and Steve Van Tuyl, who are able to follow up with changes to processes or questions that arise. An unavoidable risk with third-party analytics platforms is changing features and APIs, which may occur before, during, or after the build and will be taken into consideration when building project timelines and considering long-term sustainability needs.

Audience and Input

The initial audience for this project is the 26 institutions using, migrating, or building with Hyrax, Hyku, and/or Avalon Media System. These institutions have an immediate need for the module being developed, which will significantly enhance the analytics offered by the platform they're already engaged with. These institutions have connected with the Samvera community to indicate their interest and adoption of a Hyrax-based platform and we intend to reach out to them through Samvera community outreach events, such as Samvera Partner Calls,

Samvera Virtual Connect, and Samvera Connect in-person meetings, as well as the Hyrax Working Group that supports core maintenance and development for Hyrax.

An equally important audience is the nearly 2,000 organizations using CONTENTdm and 500 using bepress' Digital Commons internationally. A small portion of these customers were surveyed by the Hydra-in-a-Box (now Hyku) Project in 2015, some 248 institutions with 303 unique repositories.¹³ Of these, 75 or 24.7% used a CONTENTdm or Digital Commons repository. Roughly half were interested in migrating to a new platform and 17% were specifically interested in moving to Hydra-in-a-Box. Echoing the results of the HykuDirect Pilot Gap Assessment,¹⁴ reporting and object-level analytics were seen as must-haves for an ideal repository system. The analytics module built by this project will directly address gaps in the current system and speak to this audience. We intend to reach out to them through conference attendance at Open Repositories and Code4Lib, both of which take place early in the project timeline and provide opportunities for their input to shape development.

Diversity Plan

We have allocated time from user experience experts to consult at all stages of the design and build, as well as perform testing that will focus on accessibility for all users interacting with the system, who may have a variety of abilities, whether repository managers, content creators or depositors, or end users. The interface will meet Web Content Accessibility Guidelines (WCAG) 2.1. As part of the open communication and documentation plan for this project, we will contribute to ongoing conversations about how analytics can help identify web accessibility issues.¹⁵ We will also engage in community outreach throughout the project timeline and after its conclusion to ensure that any changes made to the Hyrax Analytics Working Group Requirements Tables continue to represent the needs of the community.

National Impact

Samvera has [33 active Partners](#) who use and contribute to the software. There are 60 known [in-production](#) and [in-development](#) Samvera implementations, with 23 known to be using, migrating, or building with Hyrax or Hyku and 9 using, migrating, or building with [Avalon Media System](#). With over 2,000 organizations worldwide using CONTENTdm and over 500 using bepress' Digital Commons, there is a significant potential audience for Hyrax-based platforms as open source solutions for institutional and cultural heritage repository needs and the work of the Bridge2Hyku grant will facilitate migration.

As long-term collaborators, UO and OSULP are committed to managing the ongoing maintenance needs of the analytics module as part of their ongoing Oregon Digital collaboration and contribution to the Samvera and Hyrax community. We intend for this module to be part of the Hyrax codebase and sustained over the long-term by the community, of which we will remain an active participant and partner. We will continue to share our findings after the grant is over and are interested in further exploring the intersections of analytics, usability, and accessibility.

Powerful and easy-to-use analytics make storytelling simpler. The reporting and visualizations we will build in this project can help repository managers and depositors alike demonstrate national and international impact at a glance. Analytics can help justify initial and continued investment in digital platforms that ultimately increase public access to research, data, and rare and special collections, as well as the importance of open access to scholarly and cultural heritage content.

¹³ Hydra in a Box Design Documents: <https://wiki.duraspace.org/display/samvera/Hydra-in-a-Box+Design+Documents>

¹⁴ HykuDirect Pilot Gap Assessment: <https://docs.google.com/document/d/1mpYLS2pqwKDCCxF0Kb-JTC4OjPLgmbJ6d1IBunxHqAw/edit#heading=h.up2kw3izu6dy>

¹⁵ Using Web Analytics for Mobile Interface Development: <https://doi.org/10.1080/19322909.2013.835218>

Third-party vendors continue to consolidate and strengthen their hold on additional segments of the scholarly communication and research data lifecycle. This has and will continue to impact libraries and their users ethically, financially, and in ways we cannot yet anticipate. As these concerns drive many libraries to search for alternatives, the brightest spots on the landscape are community-led initiatives like Hyrax, Avalon, and Hyku. However, many institutions cannot afford to build and manage local instances of these open source platforms. For these institutions, the possibility of hosted versions would provide a welcome alternative to third-parties with ever-increasing costs and potentially contradictory values. It is vital for these open source platforms to serve the community at their immediate point of need and to offer viable, competitive alternatives. By creating new and vital functionality for these open-source platforms, this project helps address a critical need that can help power the next generation of institutional and cultural heritage repositories across a diverse array of institutions.

Schedule of Completion	Phase 1			Phase 2			Phase 3			Phase 4			
	2019			2020									
	September	October	November	December	January	February	March	April	May	June	July	August	
Review and update the needs assessment created by the Hyrax Analytics Working Group	Project Team												
Complete institutional Request for Proposal process and hire the contractor who will be responsible for developing the analytics modules	Project Team												
Refine analytics design based on initial mockups and data from selected application(s)				Project Team & Contractor									
Coordinate requirements with the Samvera community through the Hyrax Working Group				Project Team & Contractor									
Develop the analytics modules							Project Team & Contractor						
Implement front-end visualization components							Project Team & Contractor						
Engage in usability testing							Project Team						
Engage in community outreach online	Project Team												
Engage and connect in-person at identified conferences					Project Team						Project Team		
Prepare project report evaluating the success of the project comparing initial needs assessment to the feature set ultimately available in the analytics module and by the level of adoption and feedback from the community										Project Team			



DIGITAL PRODUCT FORM

Introduction

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (e.g., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. 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

All applications must include a Digital Product Form.

- Please check here if you have reviewed Parts I, II, III, and IV below and you have determined that your proposal does NOT involve the creation of digital products (i.e., digital content, resources, assets, software, or datasets). You must still submit this Digital Product Form with your proposal even if you check this box, because this Digital Product Form is a Required Document.

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

Part I: Intellectual Property Rights and Permissions

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

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.

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

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

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

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

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

B. 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 of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

C. Metadata

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

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).

D. Access and Use

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

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

Part III. Projects Developing Software

A. General Information

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

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

B. Technical Information

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

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

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

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

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

C. Access and Use

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

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

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

Name of publicly accessible source code repository:

URL:

Part IV: Projects Creating Datasets

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

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

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

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

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

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

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

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

Name of repository:

URL:

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