

Assessment of Samvera/Hyrax for CSU Digital Collections

A Report

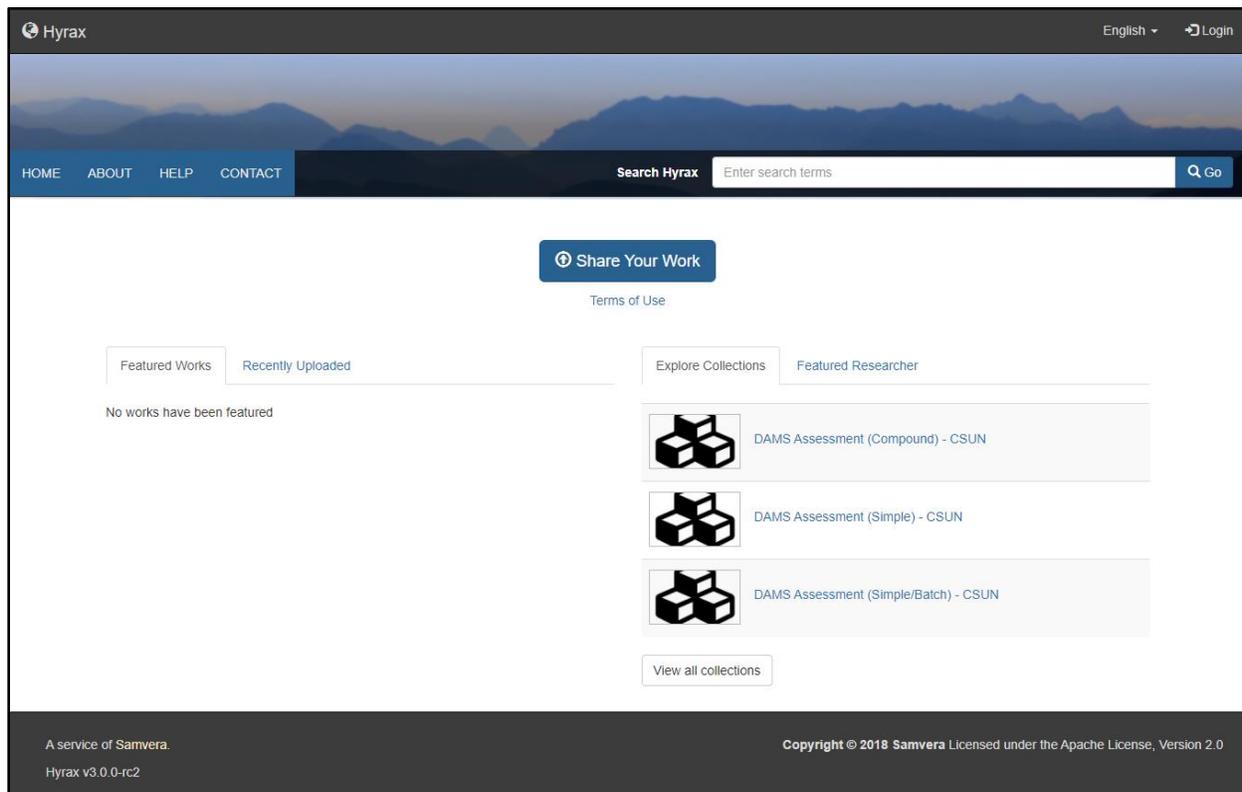


Image: Home page from CSU Sam Marcos Hyrax installation.

Prepared by the Digital Archives Working Group

March 30, 2022

Steve Kutay, California State University, Northridge

Contents

Acknowledgements.....	6
Executive Summary.....	7
Objectives	7
Results Summary.....	7
Backend Evaluation.....	7
Frontend Evaluation.....	8
Conclusion.....	9
Methods.....	10
Rubric	10
Digital Repository Assessment Tool.....	10
Referenced Resources	11
Backend.....	11
Frontend.....	11
Limitations	12
Resources.....	12
Independence	12
Objectivity	12
Costs.....	12
Results - Hyrax Backend.....	13
User Management	13
Roles (R)	13
Authentication (R).....	13
Content Types	13
Format Agnostic (R)	13
File Swapping (P).....	14
Streaming (R)	14
Object Handling	14
Derivatives (P).....	14
Object Rights Management (P).....	14
OCR/Full text (R)	15
Permanent URL (P).....	15
Embargo (P).....	15

Modeling	15
Data modeling (R)	15
Metadata.....	16
Common Schemas (R)	16
MODS (P).....	16
Controlled Vocabularies (R)	16
Linked Data (P)	16
Remediation/editing (R).....	17
Find & Replace (R).....	17
Ingest.....	17
Bulk ingest, simple (R).....	17
Bulk ingest, complex/compound (P).....	17
Add object interface (P)	18
Workflow moderation (R)	18
Quality Control (P)	18
Exports	18
Full metadata export (R)	18
Selected metadata export (R)	19
Multi format metadata extraction (R)	19
Technical metadata extraction (P).....	19
Bulk file export.....	19
Back-up/Sustainability	20
Redundancy (R)	20
Recoverability (R).....	20
Fault Tolerance	20
Network storage	20
External System Integration (Interoperability)	20
OAI-PMH (R).....	20
ULMS harvest (P).....	21
API Support (R).....	21
Reports.....	21
Analytics (R)	21
Technical/Administration (R)	21

Preservation Management	22
File Redundancy (P).....	22
XML/METS (R)	22
Fixity (P).....	22
Validation (P).....	22
Format Obsolescence (P)	22
Born-digital ingest (P).....	22
Results – Samvera/Hyrax Frontend	24
Discovery.....	24
Advanced search (R).....	24
Title, keyword, author, search (R).....	24
Customizable filters (R)	24
Delivery Interface.....	25
Zooming, Downloading, Printing (R)	25
Advanced Analysis (P)	25
Display sizing, rewind, fast forward (R).....	25
Web Pages.....	25
Displaying Lists (R).....	25
Textual Descriptions (R)	26
Campus Branding (P).....	26
Web Standards (R)	26
Accessibility (R)	26
Image Rendering/Object Embedding (R)	27
Flip Reader - Book/Newspaper Readers) (P).....	27
Exhibits (P)	27
Downloads	28
Resolution Downloads (P).....	28
Printing.....	28
User Tags/Annotations	28
Annotate	29
Social Media Support	29
Findings and Conclusion	30

Acknowledgements

This assessment was made possible by generous contributions from members of the CSU community:

- Khuong Vu, CSUSM
- Ian Chan, CSUSM
- Carmen Mitchell, CSUSM
- Azalea Camacho, CSULA
- Erick Beck, CSUS
- Alyssa Loera, CPP
- Kevin Cloud (Formerly CO)
- David Walker, CO
- Digital Archives Working Group

Executive Summary

Objectives

This report is the result of a request made by the CSU Digital Repositories Steering Committee and provides an evaluation of Samvera/Hyrax against the [*CSU Libraries DAMS Report: Requirements for Implementing Shared Digital Library Services*](#) prepared by the Digital Archives Interest Group in 2019. A rubric designed to evaluate a CSU shared digital repository platform for digital collections was deployed in order to record the level with which Samvera/Hyrax is commensurate with the published functional requirements of the 2019 report (see Methods).

This report is NOT intended to persuade or dissuade CSU campus personnel in their consideration of committing to a shared CSU DAMS as their primary repository for serving digital archival materials. The Digital Archives Working Group supports campuses in using the DAMS of their choice. We do, however, hope that users of DAMS external to a CSU shared repository will consider allowing harvests of their collections metadata in order to unify searching digital archives across all campuses of the CSU system.

Results Summary

This evaluation of Samvera/Hyrax for CSU digital collections identified 42 backend and 19 frontend required or preferred functions as identified in the 2019 [*CSU Libraries DAMS Report: Requirements for Implementing Shared Digital Library Services*](#).

Overall, this evaluation found that eight of the 61 required and preferred features were designated “not functional” (13%). One required feature was designated functional, but “does not meet requirement” (2%), and two features are “under development” (3%).

The DAMS backend assessment confirms that Samvera/Hyrax meets or exceeds 90% of the required (R) functions, and meets or exceeds 86% of the preferred (P) functions.

The DAMS frontend assessment confirms that Samvera/Hyrax meets, exceeds, or is developing 90% of the required (R) functions, and meets or exceeds 67% of the preferred (P) functions.

Backend Evaluation

REQUIRED features designated “Not functional”:

- As of 2020, no support for multi-format metadata export. However at least one source designates that development by Notch 8 is prioritized or already underway.
- No support for XML/METS metadata exports. This feature could be developed on a case need basis. There is no alternative recommendation.

PREFERRED features that are designated “Not functional”:

- No support for automatic (system-monitored) identification of obsolescent formats. Alternative recommendation: Use the manual alternative to query Fedora for specific formats in order to identify obsolescent formats to be swapped from within the system.

- No support for object rights management feature for embedding branding or watermarks into objects. Alternative recommendation is to prepare access files using common graphic software in advance of ingest.

PREFERRED features that are partially functional but designated “Does not meet requirements”:

- Incomplete support of bulk ingests of compound (complex) objects. Associated component files as part of a compound object can be pulled into the system in bulk form. However, multiple compound objects cannot be ingested together as part of a bulk compound ingest at this time. This feature could be prioritized for development on a case need basis. There is no alternative recommendation.

Notable strengths of the backend feature set:

- Excellent support for multiple user roles, which can accommodate multiple workflows and quality control operations to meet differing needs of CSU campuses.
- As a system built on multiple technologies, features can be developed (or adopted) as needed, rather than migrate all collections to another third-party system to merely address one performance need.
- The ability to assign items to multiple collections and nest collections within collections. This creates new possibilities for primary access points depending on need, as well as associating items across multiple campuses in a collective exhibit without impacting original access points such as provenance.

Frontend Evaluation

REQUIRED features designated “Not functional”:

- No support for independent image rendering, or object embedding, which enables URLs to be utilized to create a direct link to the object (rather than the object with the metadata record), as well as embedding an object directly into webpages. This feature could be prioritized for development based on a case need basis. There is no alternative recommendation.
- No support for direct printing. As a convenience, this feature was downgraded as “preferred” on the basis that downloading is enabled from which to print, which the author finds as an acceptable alternative.

PREFERRED features that are designated “Not functional”:

- No support for variable resolution downloads. However, this feature could be developed based on demonstrated need. There is no alternative recommendation.
- No support for user annotation and publicly submitted resource descriptions. However, some IIF viewers offer limited annotation, which could be implemented. An alternative recommendation for acquiring user descriptions for resources online is to utilize web forms linked from metadata and routed to designated personnel for updating descriptions.

REQUIRED features under development:

- Advanced search that includes multiple search bars, selectable fields and Boolean operators is under development at the CSU CO.
- Features have been developed for ScholarWorks to improve accessibility of the system. Needs specific to digital collections should be assessed for development.

Notable strengths of the frontend feature set are:

- The ability to implement a growing list of advanced IIIF object viewers and media players to provide advanced analysis and viewing/listening options.
- Flip readers for books and newspapers
- Exhibit support

Conclusion

Samvera/Hyrax supports a strong majority of functional requirements and preferred features of a shared CSU repository for digital collections/archives. Despite some shortcomings this reviewer finds great value in the DAMS as structured by a stack of open source technologies. This allows for improvements to be made to individual components of the system as needed by hiring external development services, OR to adopt (plug-in) alternative open source components rather than migrate to other DAMS in search of solution to any particular problem. The flexibility and extensibility of Samvera/Hyrax makes the system favorable for serving the CSU libraries digital collections for many years to come.

CSU library stakeholders are invited to assess Samvera/Hyrax functional requirements as they relate to local workflows. Subsequent campus assessments can [download the assessment tool template](#).

Methods

Rubric

In preparation for a CSU shared digital repository, the Digital Archives Working Group created an assessment tool to evaluate conformance with the 2019 DAMS functional requirements report. Given differences in functional complexity of any single feature in a DAMS, the tool deploys three strategies within its rubric. A binary scale applies to services that are simple, that is, they are either functional or not functional, such as authentication. A trinary scale applies to features that are more nuanced and could potentially be developed beyond our minimum requirement, thus a rating of ‘Optimal - Advanced features’ could be assigned. A third, granular strategy, is applied to complex features, such as advanced search capabilities, that could be functional, but still not meet our minimum requirements. Points are given to each rating from “0” to “3” depending on which scale (strategy) is used. See figure 1.

Strategy	Scoring (points applied)			
	0	1	2	3
1 (binary)	Not functional	X	Meets requirement	X
2 (Trinary)	Not functional	X	Meets requirement	Optimal (Advanced Features)
3 (Granular)	Not functional	Does not meet requirement	Meets requirement	Optimal (Advanced Features)

Figure 1: Scoring rubric features multiple strategies depending on the complexity of the function assessed.

Digital Repository Assessment Tool

The rubric is accompanied by the Digital Repository Assessment Tool created by the Digital Archives Working Group specifically for facilitating multiple assessments of Digital Asset Management Systems under consideration or development for serving the needs of CSU Libraries. Each feature assessed in the tool requires a rating label assigned by selecting from a drop-down menu. The labels include: 1) Not functional (0 points); 2) Does not meet requirements (1 point); 3) Meets requirements (2 points), or; 4) Optimal - Advanced features (3 points). Upon selecting a rating label, the tool auto-populates points as well as calculates the difference between the number of points possible and the number of points assigned to the rating. If the difference is above 2 (i.e. the assigned rating is either ‘Not functional’, or ‘Does not meet requirements’), the cell will automatically display in red, indicating a feature that requires attention for system developers (see figure 2). By incorporating a point system into the tool, multiple evaluations of the product can be averaged among multiple evaluators from the CSU libraries.

ALMA DIGITAL	Discovery in Alma Digital Collections Manager			
	Advanced search (R)	Title, keyword, author, search (R)	Customizable filters (R)	NOTES
Rubric Strategy deployed	2	3	1	
Assessment score	Not functional	Meets requirements	Not functional	Alma collection I
Numerical Score	0	2	0	
Points possible	3	3	2	
Difference	3	1	2	

Figure 2: Example of assessed category with subcategories. Selection of assessment score auto-populates a numerical score subtracted by the points possible in order to identify areas in need of development in red.

Each feature included in the evaluation contains the definition of the functional requirement it represents through comments embedded in the feature title. Hovering over each feature title will display the requirement definition. **(R) indicates a required feature and (P) indicates a preferred feature.** The definitions used are from the DAMS functional requirement report_of 2019.

The CSU Repository Assessment Tool was selected for use in this evaluation of Samvera/Hyrax in order to provide a basis for satisfying minimum functional requirements in a DAMS used in the CSU and to identify areas in need of development. The [tool used in this assessment](#) is available for review.

Referenced Resources

The following online resources were used to conduct this assessment:

Backend

- Basic (out-of-the-box) Hyrax v3.0 installation at CSU San Marcos - <https://digitools.csusm.edu/>
Does not include full technology stack.
- Samvera Knowledge Base - <https://samvera.github.io/index.html>
- About Hyrax (for Managers) FAQ - <https://hyrax.samvera.org/about/managers/>
- Fedora - <https://duraspace.org/fedora/>

Frontend

- Basic (out-of-the-box) Hyrax v3.0 installation at CSU San Marcos - <https://digitools.csusm.edu/>
Does not include Universal Viewer.
- Spotlight - <https://demo.projectblacklight.org/>
- CSU ScholarWorks - <https://scholarworks.calstate.edu/>
- Spotlight at Stanford - <https://exhibits.stanford.edu/>
- Universal Viewer - <https://universalviewer.io/>

Limitations

Limitations to this evaluation of Samvera/Hyrax for CSU Library digital collections are as follows:

Resources

This evaluation was conducted by a single CSU library representative serving in the Digital Archives Working Group as a representative of one of the CSU campuses. As such, time and commitment to the evaluation were limited, given the continued responsibilities of the evaluator. However, the scope of this assessment is relatively straight forward in its base objective to determine how the “out-of-the-box” system compares to the requirements as expressed through CSU libraries’ [CSU Libraries DAMS Report: Requirements for Implementing Shared Digital Library Services](#).

Independence

To preserve integrity, the evaluation was conducted independently of direct input from developers and users of various open source components of the Samvera technology stack and associated alternative applications, plug-ins or “gems”. As such, gaps in this analysis exist that could otherwise be filled with direct inquiries to developers and users of Samvera and associated technologies in order to clarify ambiguous findings identified within this report.

Objectivity

Despite the application of a rubric and assessment tool to guide this evaluation, the deployment of individual rubric strategies and the subsequent scores selected by the evaluator is inherently subjective. Relative experiences of any evaluator can lead to somewhat different results.

Costs

Typically, Library and Information System service costs are negotiated. Despite the importance of costs to participating CSU campuses, the financial impact of deploying Samvera as a shared DAMS for the California State University System was not a part of this evaluation. However, the Digital Repository Committee, in conjunction with the associated working groups have flagged this as a fundamental priority requiring a separate investigation.

Results - Hyrax Backend

(R) indicates a required feature, (P) indicates a preferred feature

User Management

Roles (R)

- Functional requirement: User roles should include: 1) Super-User(or administrative equivalent) that has complete configurable access; 2) Project Administrator to provide the highest levels of database maintenance and assign users with subsequent permissions; 3) Users with levels of regulated access to specific features and collections in the system, and; 4) Public User accounts to enable personal tagging/organization of online items.
- Score: 3/3, Optimal (Advanced features)
- NOTES: Robust user roles. Uses Cancancan, Ruby gem, as authorization library. Provided gems allow for the creation of more roles. Can modify user lists through YAML by default, or adding Rolify or Hydra-LDAP.

Sources:

<https://hyrax.samvera.org/about/managers/#q19>.

See Roles at <https://samvera.github.io/glossary-2.1.html>.

<https://samvera.github.io/groups.html>

<https://samvera.github.io/admin-users.html>.

See also "Setting Type Participants" at <https://samvera.github.io/collection-type-participants.html>

Authentication (R)

- Functional requirement: While the vast majority of collections will be freely accessible to the public, the system should leverage authentication technologies to restrict applicable collections to an institution's user base as needed. The application of effective authentication technologies, such as Shibboleth, should be used to manage access across variable user groups, thus expanding the overall flexibility of the DAMS to serve a variety of educational purposes otherwise confined by restrictive access relating to embargoes, copyright, etc.
- Score: 2/2, Meets requirement
- Integrates with campus authorizations such as Shibboleth, CAS.

Sources:

<https://samvera.github.io/campus-auth-integrating.html>

<https://hyrax.samvera.org/about/managers/#q19>

Content Types

Format Agnostic (R)

- Functional requirement: Both born-digital and digitized content will be migrated to IR and digital collections environments. Therefore, the DAMS must be format-agnostic to ensure all file types can be managed as needed.
- Score: 2/2, Meets requirement
- Fedora can store all files of any size.

Source:

<https://duraspace.org/fedora/resources/technical-specifications/>

File Swapping (P)

- Functional preference: In the event an ingested file must be changed, the system should accommodate the swapping of new versions of files.
- Score: 2/2, Meets preferred feature
- New versions of works can be swapped at the item display level using a drop menu (Actions) and selecting (Versions).

Sources:

<https://digitools.csusm.edu/concern/works/5138jd84x?locale=en>

<https://samvera.github.io/edit-works-2.1.html>

Streaming (R)

- Functional requirement: System should effectively serve streaming media content of audio and video formats.
- Score: 2/3, Meets requirement
- Universal Viewer (IIIF) supports streaming media delivery. Other open source viewers can be swapped.

Source:

<https://universalviewer.io/>

Object Handling

Derivatives (P)

- Functional preference: The processing of preservation quality formats as access derivatives, such as TIFFs to JPGs, will provide an optional workflow for digital preservation of analog materials as well as valuable automation.
- Score: 2/3, Meets preferred feature
- Uses gem, Derivatives to create file derivatives of uploaded content, such as access files and thumbnails from preservation files. Advanced features not determined at the time of this evaluation.

Source:

https://samvera.github.io/our_technology_stack.html

Object Rights Management (P)

- Functional preference: Object level rights management and branding should be enabled through the automatic processing of watermarks and/or embedded banners at the time of ingest. (P)
- Score: 0/3, Preferred feature not functional
- Alternatively, object rights branding is possible with external workflows using common graphics software prior to ingest.

OCR/Full text (R)

- Functional requirement: The system should provide automated OCR at ingest for full-text searching of applicable textual materials with typeface.
- Score: 2/2, Meets requirement
- Uses Ghostscript that incorporates OCR libraries from Tesseract and Leptonica.

Sources:

<https://github.com/tesseract-ocr/tesseract>
https://github.com/samvera-labs/newspaper_work
<https://ghostscript.com/blog/ocr.html>
<https://ghostscript.com/index.html>.

Permanent URL (P)

- Functional preference: Each item record is associated with a fixed and permanent universal resource locator (URL) to ensure external links to an item record are not broken.
- Score: 2/2, Meets preferred feature
- Can create handles, or Ark PURLs if needed. Can be implemented as has been done with ScholarWorks.

Source:

Director of Digital Library Services, CO.

Embargo (P)

- Functional preference: Enables the ingest of items that are subsequently suppressed from viewing for a predetermined period of time in order to accommodate issues pertaining to copyright, or permission to publish.
- Score: 2/2, Meets preferred feature
- Can manage embargoes as well as leases.

Source:

<https://samvera.github.io/lease-embargoes-2.1.html>

Modeling

Data modeling (R)

- Functional requirement: With a single Hyrax implementation serving up to 23 campuses, a common (shared) data model ensures that all campuses agree upon how content is structured and understood by the system. This allows updates to work equally across all campuses rather than updating multiple Samvera instances with different code to operate with disparate data models for each campus. Using the Resource Description Framework, the model should represent classes (i.e. collections, objects, files) and associated properties (members, relationships).
- Score: 2/2, Meets requirement
- Objects utilize the Portland Common Data Model (PCDM). Can mix/match schemas.

Source:

<https://hyrax.samvera.org/about/managers/>, https://docs.google.com/drawings/d/1-NkkRPpGpZGoTimEpYTGM1uUPRaT0SamuWDITvtG_8/edit.

Metadata

Common Schemas (R)

- Functional requirement: Descriptive schemas Dublin Core, VRA Core 3.0, PBCore are commonly used for non-bibliographic materials.
- Score: 2/2, Meets requirement
- Supports simple Dublin Core by default, but can integrate others if a campus can provide a case need and maintenance support.

Source:

<https://hyrax.samvera.org/about/managers/#q19>

MODS (P)

- Functional preference: Bibliographic descriptive schemas such as MODS will add to the extensibility of the platform. Incorporate nested data.
- Score: 2/2, Meets preferred feature
- Can be configured for MODS, however doing so requires continual maintenance for future updates.

Source: <https://github.com/samvera/hyrax/wiki/Customizing-Metadata>.

Controlled Vocabularies (R)

- Functional requirement: As a best practice of resource description, the DAMS system must enable the management of controlled vocabularies derived from both standard thesauri and local controlled vocabularies.
- Score: 2/3, Meets requirement
- Uses the Questioning Authority gem that comes configured for most common vocabularies. Has the ability to create custom vocabularies, query standard vocabularies and utilize linked data access. However, methods are not intuitive and will require implementation of tools or workflows to assist most users.

Source:

https://github.com/samvera/questioning_authority/wiki/Creating-custom-controlled-vocabularies, https://repo.samvera.org/concern/generic_works/1cf7fd24-b14a-4ab3-a251-702d2c92c878?locale=en

Linked Data (P)

- Functional requirement: Linked data implementation that provides direct access to standard vocabularies will promote consistent descriptive practice between items and collections.
- Score: 2/3, Meets preferred feature
- Uses Ruby gem LDP as linked data platform for connecting with Fedora. Full features in development.

Sources:

<https://samvera.org/samvera-open-source-repository-framework/technology-stack/>
https://github.com/samvera/questioning_authority/wiki/Creating-custom-controlled-vocabularies

Remediation/editing (R)

- Functional requirement: The platform should offer object level editing capabilities to facilitate metadata updates to ingested materials.
- Score: 2/3, Meets requirement
- Metadata can be remediated at the object and collections levels.

Source:

<https://samvera.github.io/edit-works-2.1.html>

Find & Replace (R)

- Functional requirement: The platform should offer global metadata editing capabilities within specific fields, such as searching and replacing existing values within a single field, as well as blanket changes to entire field values, such as rights statements.
- Score: 2/3, Meets requirement
- Uses Bulkcrax as a metadata editor, scheduler, and reviewer. Contains an admin interface. Scholarworks has also made implementations.

Sources:

<https://kandi.openweaver.com/ruby/samvera-labs/bulkcrax>

<https://www.youtube.com/watch?v=ykk5hQ0xSAC>

Ingest

Bulk ingest, simple (R)

- Functional requirement: The system should provide batch (or bulk) ingest of multiple items of simple objects of files and metadata.
- Score: 2/3, Meets requirement
- Uses Bulkcrax as importer, scheduler, and reviewer. Contains an admin interface. Scholarworks has also made implementations.

Sources:

<https://kandi.openweaver.com/ruby/samvera-labs/bulkcrax>

<https://www.youtube.com/watch?v=ykk5hQ0xSAC>

Bulk ingest, complex/compound (P)

- Functional preference: The system should provide batch (or bulk) ingest of multiple items of compound objects of component files and metadata.
- **Score: 1/2, Preferred feature requires development**
- Multiple files can be uploaded as separate works, and for each separate work multiple component files can be uploaded as associated with an individual work to form a compound object. However, this, by definition, is not a bulk ingest of compound objects. Bulk compound ingest could be flagged for development if a need case is demonstrated.

Source:

<https://samvera.github.io/batch-ops-2.1.html>

Add object interface (P)

- Functional requirement: The system should provide batch (or bulk) ingest of multiple items of compound objects of component files and metadata.
- Score: 2/2, Meets preferred feature
- Our testing indicates an effective add object interface and workflow.

Source:

<https://digitools.csusm.edu/>

Workflow moderation (R)

- Functional requirement: Can save and revisit projects before completing ingest or updates. (R)
- Score: 2/2, Meets requirement
- Ingest must complete, however they can remain unpublished and edited any number of times before and after publishing. You can change the structure of a work from simple to compound by associating new files with the work after ingest.

Source:

<https://samvera.github.io/edit-works-2.1.html>

Quality Control (P)

- Functional preference: The system should provide quality control functions to administrators for approving, removing, or updating submitted objects within a queue prior to the final stage of ingest, or the building of a collection's index.
- Score: 2/3, Meets preferred feature
- Uses mediated deposit to provide approval process. More complicated workflows are enabled via programming.

Source:

https://samvera.github.io/workflow_and_mediated_deposit.html

Exports

Full metadata export (R)

- Functional requirement: A collection's full metadata should be exportable for external remediation, preservation, or as downloadable files for public use such as research and machine learning operations.
- Score: 2/2, Meets requirement
- Uses Bulkcrax as exporter. Contains an administrative interface. Scholarworks has also made implementations from which to model.

Sources:

<https://kandi.openweaver.com/ruby/samvera-labs/bulkcrax>

<https://youtu.be/yZozZStmaho>.

Selected metadata export (R)

- Functional requirement: Results from a search of a collection's contents should be exportable for external remediation, or entry into other collections within the system.
- Score: 2/2, Meet requirement
- Uses Bulkcrax as exporter. Can pull selected metadata for remediation and import back into the system. ScholarWorks has also made implementations from which to model.

Source:

<https://kandi.openweaver.com/ruby/samvera-labs/bulkcrax>

<https://youtu.be/yZozZStmaho>

Multi format metadata extraction (R)

- Functional requirement: Exports should be offered in multiple formats such as delimited text, XML schemas, and METS to provide flexibility when working with databases, or for long-term preservation.
- Score: 0/3, Does not meet requirement
- Uses Bulkcrax. Supports export of common CSV format by default as of 2020. More format support is prioritized through Notch 8 developers. However, progress has not been confirmed at this time.

Sources:

<https://kandi.openweaver.com/ruby/samvera-labs/bulkcrax>

<https://youtu.be/yZozZStmaho>

Technical metadata extraction (P)

- Functional preference: System exports should include the extraction of technical metadata of items within the database.
- Score: 2/3, Meets preferred feature
- Uses File Information Tool Set (FITS) upon file ingest to extract technical metadata.

Source:

<https://hyrax.samvera.org/about/managers/>

Bulk file export

- Functional preference: System enables the export of files at the collection level, campus level, or results from a search.
- Score: 2/3, Meets preferred feature
- Uses Bulkcrax as exporter. Can export metadata from importer, collection, or work type.

Source:

<https://kandi.openweaver.com/ruby/samvera-labs/bulkcrax>

<https://youtu.be/yZozZStmaho>.

Back-up/Sustainability

Redundancy (R)

- Functional requirement: Redundancy of files and metadata to protect against catastrophic loss. (R)
- Score: 2/2, Meets requirement
- Ingest copies created in Amazon Glacier. Another copy via AWS. Campuses may wish to maintain additional copies.

Recoverability (R)

- Functional requirement: Databases must be fully recoverable in the event of catastrophic failure.
- Score: 2/2, Meets requirement
- The system as a technology stack has an advantage of limiting failures to any one specific technology (functional area) rather than rendering the entire system nonfunctional (see Fault tolerance). In the event of a catastrophic failure Fedora supports fixity checks (SHA-1, SHA-256, MD5) for single files and metadata and will compare against existing fixity digests. In addition, Glacier provides back-ups and fixity checks on ingest.

Source:

<https://wiki.lyrasis.org/display/FEDORA50/Fixity+Checking>

Fault Tolerance

- Functional requirement: System should be fault tolerant to continue operation during erroneous or compromised performance. (P)
- Score: 2/3, Meets preferred feature
- Components exist on different servers which maintains the system in use when one process goes down. Only the service is impacted. Uses separate file server.

Source:

Director, Digital Library Services, CO.

Network storage

- Functional requirement: Can Connect to external storage environments within a network such as local drives, cloud/web services, etc. typically to complete ingest or swap files. (P)
- Score: 2/2, Meets requirement
- Can connect to cloud services for ingest of existing files. Unclear whether it can store preservation files or derivatives to external storage such as AWS.

Source:

<https://samvera.org/samvera-flexible-extensible/>

External System Integration (Interoperability)

OAI-PMH (R)

- Functional requirement: The system must enable the crawling of a collection's metadata and URLs as desired via OAI-PMH to be used within an external discovery systems such as Calisphere and the Digital Public Library of America.(R)
- Score: 2/2, Meets requirement

- Currently Implemented. See Blacklight OAI-PMH. Partial with Fedora bag support. Does not fully support harvesting out of the box, but can be configured with other applications. Will provide imports via OAI-MPH.

Sources:

<https://hyrax.samvera.org/about/managers/#q19>

https://github.com/projectblacklight/blacklight_oai_provider

ULMS harvest (P)

- Functional preference: Records within the system should be harvestable by the CSU Unified Library Management System (Ex Libris Alma) via an external API or another protocol.
- Score: 2/2, Meets preferred feature
- Blacklight OAI Provider plugin can be implemented. A customization solution already provided for ScholarWorks.

Source:

<https://hyrax.samvera.org/about/managers/#q19>

API Support (R)

- Functional requirement: Support available to campus personnel that create customizations. (R)
- Score: 2/2, Meets requirement
- Generally, CO personnel will provide API support. Will consider allowing other campus personnel contribute support where applicable. Uses customizable HTTP APIs to communicate with Fedora (storage) and Solr (index).

Source:

https://samvera.github.io/our_technology_stack.html

Reports

Analytics (R)

- Functional requirement: Collection-level and item-level usage statistics and analytic reports will provide valuable data to enable administrators to strategize development and promotion of their collections.
- Score: 2/3, Meets requirement
- Multiple stats available via dashboard. Leverages Google analytics. Limited, but new module coming soon, according to DLS Director.

Source:

<https://hyrax.samvera.org/about/managers/#q19>

Director of Digital Library Services, CO

Technical/Administration (R)

- Functional requirement: Reports such as items added, file types, objects per collection, users, etc.
- Score: 2/3, Meets requirement
- Our testing indicates an acceptable variety of reports to cover users, collections and objects.

Source:

https://digitools.csusm.edu/catalog?f%5Bdepositor_ssim%5D%5B%5D=stephen.kutay%40csun.edu&f%5Bgeneric_type_sim%5D%5B%5D=Collection&locale=en

Preservation Management

File Redundancy (P)

- Functional preference: Redundant file storage of preservation files
- Score: 2/2, Meets preferred feature
- Ingest copies of preservation files in Amazon Glacier.

XML/METS (R)

- Functional requirement: Structured metadata exports in XML/METS (see Exports)
- **Score: 0/2, Not functional**
- Could develop if a case need is demonstrated.

Source:

Director of Digital Library Services, CO

Fixity (P)

- Functional preference: Integrity checksums (e.g. MD5, SHA256)
- Score: 2/2, Meets requirement
- Fedora supports fixity checks (SHA-1, SHA-256, MD5) for single files. Will compare against existing fixity digests. Glacier provides fixity checks on ingest.

Source:

<https://wiki.lyrasis.org/display/FEDORA50/Fixity+Checking>

Validation (P)

- Functional preference: Format validations (e.g. JHOVE)
- Score: 2/2, Meets requirement
- Uses FITS through Hyrax File Characterization. JHove and DROID included in FITS. Extracts technical metadata.

Format Obsolescence (P)

- Functional preference: Automated reformatting of obsolescent formats (P)
- **Score: 0/2, Preferred feature not functional**
- Obsolescence monitoring or automated format migrations are not confirmed and unlikely. Fedora can be queried for file formats in order to inform preservation management decisions, which is accomplished via file swaps.

Born-digital ingest (P)

- Functional preference: Ingest of born-digital materials for preservation
- Score: 2/3, Meets preferred feature
- System is format agnostic and capable of ingesting all file types (including files/data created within a born-digital domain). However, the system does not serve as a standalone preservation

management system. Preservation priorities and workflows for migrating and managing preservation files and metadata (including from fixed storage systems) are the responsibility of each campus. [Born-digital preservation management guidelines and resources](#) have been created by the Digital Archives Working Group for voluntary use by CSU campuses.

Results – Samvera/Hyrax Frontend

(R) indicates a required feature, (P) indicates a preferred feature

As a technology stack, Hyrax uses Blacklight and Spotlight, and is able to implement a variety of IIF viewers and audiovisual media playback engines to enable user searching and delivery of materials within a digital asset management system. The assessment of specific functional requirements for the DAMS frontend are below.

Discovery

Advanced search (R)

- Functional requirement: Search engine should provide advanced search (multiple Boolean and specified field search bars). Search across specific fields.
- **Score: Under development**
- Hyrax v3 out of the box, does not offer multiple search bars using boolean operators and variable field assignments. However, the CO is planning to build this feature based on the ScholarArchive@OSU.

Source:

Director of Digital Library Services, CO

Title, keyword, author, search (R)

- Functional requirement: Search engine should provide, specifically, keyword, author and title search capabilities.
- Score: 2/3, Meets requirement
- Search box offers field specific searching for 'All items', 'My Works', and 'My Collections'. Other specific values must be added.

Source:

https://digitools.csusm.edu/catalog?utf8=%E2%9C%93&locale=en&view=gallery&search_field=all_fields&q=iolanthe

Customizable filters (R)

- Functional requirement: Discovery layer should provide customizable filters and facets. (R)
- Score: 3/3, Optimal (Advanced features)
- Uses Blacklight, which truncates long facets, but enables popups for complete lists. Contains graphic display of time distribution of documents. Out of the Box Hyrax V3 utilizes facets.

Source:

https://digitools.csusm.edu/catalog?utf8=%E2%9C%93&locale=en&view=gallery&search_field=all_fields&q=iolanthe, <https://samvera.github.io/blacklight-plugins.html>
<http://projectblacklight.org/>

Delivery Interface

Zooming, Downloading, Printing (R)

- Functional requirement: PDF viewers should be enabled and effective regarding zooming, downloading, and printing.
- Score: 2/3, Meets requirements
- Offers download from which to print.

Source:

<https://digitools.csusm.edu/concern/works/8k71nh08w?locale=en>

Advanced Analysis (P)

- Functional preference: IIIF and other embedded viewers will facilitate advanced analysis of image files.
- Score: 3/3, Optimal (Advanced features)
- Contains gem RIIIF to install the IIIF universal viewer (<https://universalviewer.io/>) More advanced viewers, such as Mirador (<https://projectmirador.org/>) and OpenSeadragon (<https://openseadragon.github.io/>) can be installed by consulting the work of other institutions (<https://openseadragon.github.io/examples/in-the-wild/>).

Source:

https://samvera.github.io/troubleshooting_riiif.html

Display sizing, rewind, fast forward (R)

- Functional requirement: Audiovisual media players should enable basic functions such as display sizing, rewind and fast forward.
- Score: 3/3, Optimal (Advanced features)
- Uses IIIF Universal Viewer for image display and audiovisual playback. Other IIIF viewers can be implemented.

Source:

<https://universalviewer.io/>

Web Pages

Displaying Lists (R)

- Functional requirement: The system should provide the means for displaying lists of collections (e.g. splash pages) as well as the means for displaying these according to each participating campus. This could be a native function of the software and/or a customizable web-based option.
- Score: 3/3, Optimal (Advanced features)
- Can display collections across all campuses assuming all CSU campuses share the same instance. Since collections can be nested and items can exist in multiple collections, exhibits can use items from multiple campuses. In addition, metadata will identify contributing campuses therefore, the built-in facet system can be used to filter campus collections.

Source:

https://digitools.csusm.edu/catalog?f%5Bhuman_readable_type_sim%5D%5B%5D=Collection&locale=en
<https://samvera.github.io/configuration-2.1.html>

Textual Descriptions (R)

- Functional requirement: The system should provide the means for context through textual descriptions (e.g. landing pages) of the collections in order to give users critical context associated with the materials contained therein. This could be a native function of the software and/or a custom web-based option, such as a WYSIWYG editor.
- Score: 3/3, Optimal (Advanced features)
- Collection descriptions are supported. Advanced descriptions are enabled by the plugin, Spotlight. Furthermore, collections nested within a collection are supported. This has the potential for shared campus exhibits of archival collections of different provenances across the CSU.

Source:

<https://digitools.csusm.edu/collections/kh04dp681?locale=en&view=gallery>
<https://spotlight.projectblacklight.org/>

Campus Branding (P)

- Functional preference: The system should provide the means for campus branding of the collections through graphics, wordmarks, and color assignments
- Score: 2/3, Meets preferred feature
- Basic branding is supported. Colors and logos are available in Hyrax and enabled by the CO.

Source:

<https://digitools.csusm.edu/dashboard/collections/kh04dp681/edit?locale=en#branding>

Web Standards (R)

- Functional requirement: The user interface must be responsive according to web standards.
- Score: 2/2, Meets requirement
- Basic branding is supported. Colors and logos are available in Hyrax and enabled by the CO. Fedora complied with modern web standards.

Source:

<https://duraspace.org/fedora/resources/technical-specifications/>

Accessibility (R)

- Functional requirement: The system must be compliant with the Americans with Disabilities Act (ADA) to ensure accessibility, or is committed to reaching full compliance.
- **Score: Under development**
- Blacklight accessibility is likely mixed. There are no known (public) comprehensive tests of ADA compliance. However, the Blacklight plugin, GeoBlacklight, claims to fulfill WCAG 2.0, Section

508, and other accessible standards. Beyond this, there is at least one claim in the library community that Blacklight has performed reasonably well among accessibility tests, with some development still required.

Source:

<https://duraspace.org/fedora/resources/technical-specifications/>
<https://github.com/projectblacklight/spotlight/issues/2106>
<https://geoblacklight.org/guides.html>

Accessibility developments have been made for ScholarWorks. Hyrax contains task forces to address accessibility issues. Will need to consider accessibility development locally based on needs for a shared digital collections DAMS.

Source:

Director of Digital Library Services, CO

Image Rendering/Object Embedding (R)

- Functional requirement: With each item created, the system should be able to associate that item with a direct link to the fully rendered object file that can then be reused in third party tools (e.g. exhibit platforms, webpages) for image rendering and object embedding using a stable URL.
- **Score: 0/3, Not functional**
- Service not currently available, but can be customized on demonstrated case need.

Source:

Director of Digital Library Services, CO

Flip Reader - Book/Newspaper Readers) (P)

- Functional preference: Special object display such as side-by-side page flip view, or news article highlight/selection.
- Score: 3/3, Optimal (Advanced features)
- Uses Newspaper_Works, a rails gem to serve newspaper content. Contains both back and front-end feature sets.

Source:

https://github.com/samvera-labs/newspaper_works/wiki/Features-List

Exhibits (P)

- Functional preference: As an identified priority of the survey respondents, it is requested that the system provide the means for curating exhibits online. One complicating factor is that no mutually agreed upon definition of what constitutes an exhibit online was achieved. Some digital asset management systems offer the means to create web pages, however limited, toward an exhibit-like function. In the absence of a dedicated exhibit function, the system should utilize plug-ins, extensions, or integrated applications designed for this purpose. (P)

- Score: 3/3, Optimal (Advanced features)
- Uses Spotlight to create feature rich websites for curating digital collections materials. As an open source technology, newly developed widgets, etc. can be implemented, making the platform extensible moving forward.

Source:

<https://spotlight.projectblacklight.org/>

Downloads

Resolution Downloads (P)

- Functional preference: The system should provide full and partial resolution downloads of the materials within the database to enable flexible use in research and course learning objectives. (P)
- **Score: 0/3, Preferred feature not functional**
- The system does not offer the means to select and download varying resolutions of images. Downloads are enabled at full resolution. Customization of this service could be requested on demonstrated need.

Source:

<https://digitools.csusm.edu/concern/works/5138jd84x?locale=en>

Printing

- Functional requirement: The system should enable printing of reasonably of file types within the database. (R)
- **Score: 0/3, Does not meet requirement**
- Only downloads are offered, but as an alternative, the item can be printed upon download. Given this knowledge, direct printing is a convenience, and therefore this feature has been downgraded to “Preferred” (P).

Source:

<https://digitools.csusm.edu/concern/works/8k71nh08w?locale=en>

User Tags/Annotations

Tag items

- Functional requirement: User functionality is greatly improved by the ability for the public to tag individual items for future use. (P)
- Score: 2/3, Meets preferred feature
- Items can be bookmarked in the gallery view.

Source:

<https://digitools.csusm.edu/collections/kh04dp681?locale=en&view=gallery>

Annotate

- Functional requirement: Users and custodians of archival materials benefit significantly from the ability of users to annotate. (i.e. provide public descriptions) of the database items therein. (P)
- Score: 0/3, Does not meet preferred feature.
- No ability to annotate objects within the system. However, some IIF viewers can accommodate certain annotations. Also, the system can provide links to web forms in order to submit public descriptions to specific personnel.

Social Media Support

- Functional requirement: The system should enable sharing through social media by building in links to popular social media sites accessible at the item level through the use of stable URLs and embed codes, or in the least, provide easy access to stable URLs and embed codes. (P)
- Score: 2/3, Meets requirement
- Facebook, Twitter and Tumbler are enabled out of the box at the work level, new social media links can be customized.

Source:

<https://digitools.csusm.edu/concern/works/5138jd84x?locale=en>

Findings and Conclusion

See Executive Summary, page 7.

Send questions or comments to stephen.kutay@csun.edu.