

# Final Report

## University of Houston Libraries Scholarly Works Pilot Project

Expanding Access to UH Libraries Scholarly Works

January 2017

Anne Washington, Metadata Librarian, Project Lead

Adam Townes, Coordinator of Research Support Services, Copyright/Research

Andy Weidner, Metadata Services Coordinator, Project Technical Advisor

Santi Thompson, Head of Digital Repository Services, Project Champion

Annie Wu, Head of Metadata and Digitization Services, Project Champion

# Table of Contents

<b>Executive Summary</b>	<b>2</b>
<b>Introduction</b>	<b>3</b>
<b>Background</b>	<b>3</b>
Submission Model	3
Types of Work	4
<b>Methodology</b>	<b>5</b>
Planning Phase: March-April 2016	5
Phase 1: May-July 2016	5
Phase 2: August-October 2016	6
<b>Results</b>	<b>7</b>
Participation	7
Feedback	8
Survey	8
Focus Group	9
SWOT Analysis	9
<b>Recommendations</b>	<b>11</b>
<b>Appendix A: Pilot Project Participation Data</b>	<b>13</b>

## Executive Summary

From March to October, 2016, the University of Houston Libraries Scholarly Works Pilot Project team developed and tested processes for submitting librarian and library staff scholarly works to the University of Houston Institutional Repository. The resulting standards, workflows, and documentation have laid the foundation for UH Libraries to expand this service to the broader UH scholarly community.

The team chose to implement the mediated self-submission model instead of the fully mediated or self-submission approach as it offers flexibility for the depositor and offers an opportunity to for library staff to review and correct metadata. The team limited the scope of acceptable works to conference presentations, posters, and papers as well as journal articles where the author clearly held copyright.

The project was comprised of three phases: a planning phase, a test period with limited departments, and a final phase inviting work submissions from all librarians and staff. Overall, there were a total of 14 participants and a total of 109 works submitted. The majority of participants submitted 1 or 2 works. These depositors represented Administration, Branch Services, Digital Research Services, Liaison Services, Metadata and Digitization Services, Resource Discovery Systems, Special Collections, and Web Services - 8 of the 12 Libraries departments.

Feedback from participants indicated that this was a useful service with the repository seen as a reliable place to store items referenced on a CV. Key points of improvement identified through informal feedback, a formal survey, focus group, and a team SWOT analysis include: simplifying the submission process, expanding the types of works accepted, and increasing the visibility of the service.

The Scholarly Works Pilot Project Team recommends permanent adoption of the Institutional Repository for University of Houston faculty, librarian, and staff works. To make this a truly valuable service to the UH community, the IR requires further development. The team recommends the following next steps:

1. Continue use of the mediated self-deposit model.
2. Continue technical development to address UH Libraries pilot project feedback.
3. Broaden the scope of accepted works.
4. Consider preservation actions for Institutional Repository content.
5. Engage in outreach and education to promote repository services

Work on these recommendations will continue with the Cougar Scholar Open Access Team, a Year 1, Phase 1, SPIET Team.

## Introduction

Each year, University of Houston (UH) librarians and library staff produce a wealth of scholarly content. This body of work represents a substantial and meaningful contribution to the library profession. However, UH librarian and library staff scholarly output is currently published in a haphazard, distributed fashion, scattered across the internet on conference websites, open access journals, and LinkedIn profiles. The Scholarly Works Pilot Project developed workflows and procedures to gather these works into a centralized repository. The project has effectively expanded UH Library repository services and advanced the Libraries' goal of becoming one of the leaders in research productivity on campus. The development of standards, workflows, and documentation during the pilot project has laid the foundation for UH Libraries to expand this service to the broader UH scholarly community.

## Background

To begin work on the project, the team reviewed peer universities' institutional repositories and literature in the field to determine the submission model to employ and the types of work accepted for the pilot.

## Submission Model

There are three prevailing models for authors' submission of work into an institutional repository: mediated deposit, self-deposit, and mediated self-deposit. Mediated deposit means that the author contacts a designated depositor, sends them the file and basic metadata, and the depositor describes, then uploads the work into the institutional repository. In a self-deposit model, the author describes and deposits their work directly into the IR, with the work either becoming immediately available or available after a chosen embargo date.

The benefit of a fully mediated submission is that a librarian is able to review and describe the content according to best practices. However, this is time consuming and, depending on the number of items received, may require a significant time commitment. To reduce library resources needed to mediate deposits and to allow faculty, staff, and other depositors more flexibility and control over their work, the self-deposit model has gained in popularity. Monica Rivero, Digital Curation Coordinator at Rice University's Fondren Library, noted that faculty preferred the unmediated self-deposit model "as they can immediately see their work online and make metadata edits as needed." However, an unmediated self-deposit approach can result in inconsistent metadata. For example, as authors enter names and subjects in various ways, it becomes more difficult to collocate works by the same author or on the same subject. The mediated self-deposit model offers flexibility for the depositor in that they may deposit the works

they choose at whatever time they choose. It also provides an opportunity to for library staff to review and correct metadata as well as other errors resulting high quality repository data.

The team chose to implement a workflow using a mediated self-submission process. Authors described and deposited their work directly, then a designated team of library staff reviewed the submission, mediated error correction, and published the work. This process maximized the possibility for submission while minimizing staff mediation. The limited staff mediation helped ensure compliance and metadata accuracy, building a quality foundation for an expanded institutional repository.

## Types of Work

The types of work accepted into the institutional repositories reviewed varies in scope from very limited to very inclusive. Baylor's BEARdocs includes Electronic Theses and Dissertations (ETDs), student papers, and scholarly articles. Rice University's Digital Scholarship Archive includes a much broader scope of work. It includes faculty research, ETDs, as well as the digitized collections of rare or unique books, images, musical performances, and manuscripts. The University of North Texas' Scholarly Works collection includes both published and unpublished faculty and staff research and scholarship as well as students' professional work. University of Cincinnati's Scholar@UC repository accepts "scholarly output and creative works of UC faculty (current and emeritus), staff, and students" including grey literature, pre-, post-, and published versions of academic papers, manuscripts, senior design projects, theses and dissertations, posters, datasets, presentations, and other creative works<sup>1</sup>. Such breadth of work creates a fuller picture of the community's scholarly output, but complicates technical requirements, metadata requirements, and workflows.

The team determined a limited scope of eligible work types for the pilot project. Scholarly materials must have been presented in a public forum (e.g., conference, journal) outside of the UH Libraries while the creator was an employee of the UH Libraries. The pilot only included the following types of scholarly work, submitted in PDF format:

- Conference Presentation
- Conference Poster
- Conference Paper, Published Version
- Peer Reviewed Journal Article, Published Version
- Non-Peer Reviewed Journal Article, Published Version

To manage risk, the submissions were required to be open access or the depositor held the copyright to publish the work. Depositors were responsible for determining the copyright status of potential submissions. Within the context of the pilot project, scholarly works are those works resulting from participation in scholarly activities such as authoring books or articles, book

---

<sup>1</sup> Scholar@UC Collection Policy, [https://scholar.uc.edu/coll\\_pol\\_request](https://scholar.uc.edu/coll_pol_request)

reviews, conference presentations publications, and presenting posters. This range of accepted materials is within the stated UH Libraries collection development policy working draft for the institutional repository<sup>2</sup>.

## Methodology

The project was comprised of three phases: a planning phase, a test period with limited departments, and a final phase inviting work submissions from all librarians and staff.

### Planning Phase: March-April 2016

The goals of the planning phase were to establish formal workflows, intranet documentation, and to hold an Open Forum to introduce the project to the library. The team first worked to develop technical infrastructure, reviewing the functionality offered by DSpace and developing a list of customizations required before the submission period would begin. Part of this work included developing a metadata application profile listing the fields, requirements, and input guidelines<sup>3</sup>. DSpace customizations included adding or modifying metadata fields according to the application profile, adding help text to the submission form, and configuring the deposit and review steps in DSpace. To prepare for the Open Forum and for the first phase of submissions, the team developed and shared documentation detailing submission procedures for depositors, review procedures for staff in the Metadata Unit, and an FAQ document<sup>4</sup>.

### Phase 1: May-July 2016

The goals of Phase 1 were to solicit submissions from select UH Libraries departments, troubleshoot workflow issues, and further develop training resources. To begin, two training sessions were held for interested participants in the Metadata and Digitization Services, Web Services, and Digital Repository Services departments. These sessions included project background, a walkthrough of the submission process, and the opportunity for attendees to ask questions and give feedback (See Appendix A for Attendance Data).

Once the IR submission period officially started, DSpace functions and limitations became more apparent. The Metadata Unit was able to better understand how the review and publication processes worked and updated workflows and documentation accordingly. However, technical issues also surfaced which required significant troubleshooting efforts.

---

<sup>2</sup> Digital Collections Development Policy – Working Draft,  
[http://info.lib.uh.edu/sites/default/files/docs/UHLibrariesDigColDevPolForWeb\\_20151123.pdf](http://info.lib.uh.edu/sites/default/files/docs/UHLibrariesDigColDevPolForWeb_20151123.pdf)

<sup>3</sup> <http://intranet.lib.uh.edu/node/7636>

<sup>4</sup> <http://intranet.lib.uh.edu/scholarly-works-pilot>

When the project team first developed the submission workflow, the aim was to take advantage of ORCID<sup>5</sup> integration in DSpace. A challenging aspect of any digital library system is disambiguation of names. ORCID addresses this issue by providing researchers with unique identifiers helping to accurately link creators with their works. DSpace appeared to offer integration with ORCID by providing an ORCID database lookup function as an option for populating the Author field. The first version of submission guidelines instructed depositors to use the lookup feature to enter author names. Halfway through the first phase, feedback indicated that the use of this function negatively impacted discovery and browsing. The team worked with TDL to address this issue and ultimately, due to system limitations, had to remove the ORCID “integration” from the workflow. Including the author’s ORCID was still a priority, so the team instructed depositors to add their ORCID to the Identifier field with hopes that this data could be used in the future to make richer connections to ORCID and our local authority system.

In order to troubleshoot the ORCID issues, the team had to cut the original submission window short by two weeks. Still, in the four week submission window, we received 76 submissions from 4 depositors. The remainder of this phase was used to correct metadata affected by the ORCID issues and make other system customizations. The team provided participants with an opportunity to submit feedback informally and through a webform. This feedback informed changes including edited help text, additions to the depositor guidelines, and improved notification e-mails.

## Phase 2: August-October 2016

The goals of Phase 2 were to solicit submissions from all UH Libraries departments, provide training for depositors, refine existing workflows and resources, develop new training resources, and continue to seek feedback from participants.

In the first weeks of this phase, the team held two training sessions for interested participants throughout the UH Libraries (See Appendix A for Attendance Data). In this training session, presenters attempted to stress and clarify points that had been confusing according to Phase 1 feedback. Presenters also introduced new guidelines for name and ORCID entry. Depositors were asked to enter their name consistently across submissions and to include only their ORCID in the submission metadata. The Metadata Unit would then create an authority in our local linked data vocabulary manager, Cedar<sup>6</sup>, with the author’s name and ORCID. The vocabulary manager data can later be used for metadata remediation and future system integrations. In this phase, 10 new depositors submitted a total of 18 works.

Other activities in this phase included: continued development and documentation of the submission review workflow in the Metadata Unit, collaboration with the Resource Discovery Systems department to include works from our IR in OneSearch results, and a call for feedback both through a survey and a focus group session.

---

<sup>5</sup> <https://orcid.org/>

<sup>6</sup> <https://vocab.lib.uh.edu/en.html>

# Results

## Participation

Overall, there were a total of 14 participants and a total of 109 works submitted. The majority of participants submitted 1 or 2 works. Phase 1 included 4 submitters who deposited a total of 91 works. This group included one outlier who submitted 57 items. These depositors represented the limited departments selected for participation in this phase: Digital Research Services, Metadata and Digitization Services, and Web Services. In Phase 2, we received 18 submissions from 10 new depositors. These depositors represented 5 additional library departments: Administration, Branch Services, Liaison Services, Resource Discovery Systems, and Special Collections. See Table 1.

Department	# Participants	# Submissions
Administration	1	1
Branch Services	1	1
Digital Research Services	1	16
Liaison Services	4	4
Metadata and Digitization Services	3	20
Resource Discovery Systems	1	5
Special Collections	2	5
Web Services	1	57
<b>Total</b>	14	109

Table 1. Number of participants and submissions by department

While both staff and librarians were invited to submit works to the repository, librarians of various ranks were the only depositors. Of the 14 participants, 10 were Assistant Librarians and the others were of higher ranks: Associate Librarian, Librarian, and Dean. Some staff members are represented in the IR, as they are co-authors of works, but they did not deposit those works. See Table 2.



<b>Rank</b>	<b># Participants</b>	<b>#Submissions</b>
Assistant	10	48
Associate	2	3
Librarian	1	57
Dean	1	1
<b>Total</b>	14	109

Table 2. Number of participants and submissions by rank

See Appendix A for additional data.

### Feedback

Informal feedback was solicited throughout the entirety of the project. Additional insight was gained through a formal survey and focus group conducted at the end of Phase 2 and the Scholarly Works Pilot Project Team’s SWOT Analysis.

### Survey

Survey results indicated that most project participants felt that the repository interface was intuitive to use. This suggested that should the current interface design and organization be retained, future repository users should be able to submit content with relatively little training. This proposition is further supported by participants responding affirmatively that the submission process was easy to follow and the sequence of steps involved in the process was clear. Participants also indicated that they thought it was easy to complete the submission process.

Most survey participants responded that the overall process of submitting to the repository was simple and easy to follow. However, respondents noted a number of difficulties with different aspects of the process, or confusion about what to enter into the information fields. Participants experienced some uncertainty regarding the information that they should enter in relation to ORCID’s. Users desired more regarding whether they should submit only their own ORCID, or whether the ORCID’s for the author and all co-authors should be included. Users also expressed difficulty with finding the location of the *Start a New Submission* link. There also appears to a need for additional descriptive and instructional text for some of the text fields.

In terms of their motivations to use the repository, survey participants articulated a number of reasons. A primary reason participants cited for using the repository was the ability to include permanent links to their scholarly work in the repository on their CV’s. Participants also explained a desire to use the repository as a form of storage, and as a means of organizing and and easily accessing their file over time.

## Focus Group

Results from analysis of data collected from a focus group largely support the results from the survey feedback. The difficulties expressed by the focus group participants confirm and elaborate on those noted by survey participants. When asked what came to mind when thinking of their experience with the scholarly works repository, participants initial responses were very similar.

After a brief discussion, group members reached consensus and stated that the submission process felt long. Further, the focus group also noted that they experienced some difficulty in locating the repository on the staff intranet. Their difficulty was further compounded by the placement of the link to begin a submission. The radio button elements in the submission process were also confusing for participants and the group noted that they did not like using it. Another issue noted by the group members revolved around their difficulty with using the search function within the repository platform.

The group also voiced their frustration with the inability to utilize their ORCID's during submissions. Such frustration over ORCID's may result from the exclusive nature of the pilot, library employees have a greater familiarity with ORCID than the average faculty member or student who may not be as concerned with using ORCID's. Additional data from subsequent phases of repository implementation may invalidate this supposition, but this is doubtful given the number of ORCID holders on campus among students and faculty, which stands at around 500 for the central campus. This relatively low number, suggests either little interest, or awareness about ORCID among the UH student and faculty population.

Issues and difficulties aside, participants provided some positive feedback, identified some benefits of using the repository, and suggested offsetting the lengthy submission process. The group saw the repository as a reliable place to store items referenced on their CV's. Group participants also noted that while they thought that the help materials were useful, they suggested that they could be better integrated into the actual submission form and that field instructions could be made clearer. Participants also suggested utility of an autofill function that could auto-populate fields based on previous submissions and potentially shorten the submission process. Another useful suggestion was the provision of boilerplate examples for file and entry descriptions to help guide users through the process and to encourage consistency in entries. The group also suggested the need for a more gratifying experience once submission was complete. Such gratification might come in the form of a congratulatory email or a message thanking the user for using the repository.

## SWOT Analysis

Following completion of Phase 2, the pilot project team began assessment of project outcomes. A Strength, Weakness, Opportunity, and Threat (SWOT) analysis was determined be a useful means of assessment and review. Below is a summary of the SWOT analysis.

## **Strengths**

- The scholarly works repository will serve as a gateway to open access.
- The submission process becomes easier after the first submission and the user becomes more familiar with the process.
- The repository will make scholarly works more available, visible, and discoverable.
- The blended approach to provision of repository services between back of house and front of house outreach and instruction strengthens interdepartmental connections.
- The pilot project has served as a proof of concept for the mixed team approach to implementation of strategic plan initiatives and projects.

## **Weaknesses**

- System was not ready for hands-on use during training sessions
- Many pilot project participants indicated that submission and ingest process was far too long and complicated.
- Users indicated that the relevant documentation is not well integrated into interface and could be made clearer to ease the deposit process.
- The ingest process is currently imbalanced, particularly in regard to the number of steps required for ingest and the number of metadata fields users must fill in.

## **Opportunities**

- There is the potential to expand scope of works and files types to be accepted such as grey literature, internal presentations, and committee reports.
- Continued development of the repository and the necessary expertise to support continued operations would allow the library to become a campus leader in scholarly communication.
- Expand service offerings to support faculty promotion and tenure. Emphasize the utility of repository as a way to highlight research productivity.
- Strengthen supporting and guiding documentation with continued user feedback.
- Further refine the cross-departmental workflows and solidify interdepartmental connections and partnerships.

## **Threats**

- The limitations inherent to DSpace in regard to development, configuration, and customization represent a potential challenge to tailoring the system to UH needs.
- Upload and file size limitations are issues that have been noted by both potential users and pilot project participants.
- Potential lack of interest or resistance to buy-in for repository.
- Staffing needs for continuation of this effort requires commitment from lots of departments.

- Multiple repositories that require management, each with own characteristics and idiosyncrasies.

## Recommendations

The Scholarly Works Pilot Project Team recommends permanent adoption of the Institutional Repository for University of Houston faculty, librarian, and staff works. To make this a truly valuable service to the UH community, the IR requires further development. The team recommends the following next steps:

- 1. Continue use of the mediated self-deposit model.**
  - The team recommends a faculty pilot with strong liaison librarian involvement using the mediated self-submission process. This method allows for metadata consistency, most importantly with regard to author names, creating a basis for more discoverable resources. However, we recommend room for adjustment of this model based on faculty pilot feedback and submission volume.
- 2. Continue technical development to address UH Libraries pilot project feedback.**
  - An exploration of technical re-tooling is required to address issues identified by project participants. Primary areas of focus are adjustments to the length of the process, in-process help text improvements, and increase visibility of the repository and its supporting documentation.
- 3. Broaden the scope of accepted works.**
  - Expanding the scope of eligible participants, work types, and file types may increase participation. Undergraduate and graduate students should be considered a potential audience for this service. Copyright issues may have prevented interested users from submitting works during the pilot project. Accepting additional types of work such as pre-prints, post-prints, and grey literature, would strengthen the case for this service. This may require a review of the University of Houston Libraries' Digital Collections Development Policy working draft.
- 4. Consider preservation actions for Institutional Repository content.**
  - One benefit we may offer depositors is the digital preservation of works they submit to the institutional repository. Clarifying the extent to which IR content is preserved must be defined and communicated accurately to depositors.
- 5. Engage in outreach and education to promote repository services**
  - Librarians and staff can leverage connections with UH administration, faculty, staff, and students to increase the visibility and use of this service. In order to do this successfully, it will be minimally required to create a web presence for these services and educational materials as well as an action plan for outreach.

Work on these recommendations will continue with the Cougar Scholar Open Access Team, a Year 1, Phase 1, SPIET Team.

## Appendix A: Pilot Project Participation Data

### Phase 1:

Submission period:

Planned: 5/12/2016 - 7/8/2016 (8 weeks)

Actual: 5/24/2016 - 6/22/2016 (4 weeks)

Phase 1 participation:

Workshop 1, 5/12/2016 attendance: 6

Workshop 2, 5/19/2016 attendance: 2

Submitters: 4

Depositor	Department	Rank	Number of Submissions
Depositor 1	Web Services	Librarian	57
Depositor 2	Digital Repository Services	Assistant	16
Depositor 3	Metadata and Digitization Services	Assistant	16
Depositor 4	Metadata and Digitization Services	Associate	2
		<b>Total:</b>	<b>91</b>

### Phase 2:

Submission period:

Planned: 8/16/2016-10/14/2016 (8.5 weeks)

Actual: 8/29/2016-10/14/2016 (6.5 weeks)

Phase 2 participation

Workshop 1, 8/16/2016 attendance: 15

Workshop 2, 8/24/2016 attendance: 5

Personal appointments: 3

Submitters: 10

Name	Department	Rank	Number of Submissions
Depositor 5	Resource Discovery Systems	Assistant	5
Depositor 6	Liaison Services	Associate	1
Depositor 7	Liaison Services	Assistant	1
Depositor 8	Administration	Dean	1
Depositor 9	Branch Services	Assistant	1

Depositor 10	Liaison Services	Assistant	1
Depositor 11	Liaison Services	Assistant	1
Depositor 12	Special Collections	Assistant	2
Depositor 13	Special Collections	Assistant	3
Depositor 14	Metadata and Digitization Services	Assistant	2
		<b>Total:</b>	<b>18</b>

**Overall:**

Total participants: 14

Total submissions: 109

Department	# Participants	# Submissions
Administration	1	1
Branch Services	1	1
Digital Repository Services	1	16
Liaison Services	4	4
Metadata and Digitization Services	3	20
Resource Discovery Systems	1	5
Special Collections	2	5
Web Services	1	57
<b>Total</b>	14	109

Rank	# Participants	#Submissions
Assistant	10	48
Associate	2	3
Librarian	1	57
Dean	1	1
<b>Total</b>	14	109