

2023

## Envisioning the Future of a Mature IR: A Midlife Assessment of ScholarWorks@UMassAmherst

Erin Jerome

*University of Massachusetts Amherst, ewjerome@library.umass.edu*

Thea P. Atwood

*University of Massachusetts Amherst, tpatwood@indiana.edu*

Melanie Radik

*University of Massachusetts Amherst, mradik@umass.edu*

Rebecca M. Seifried

*University of Massachusetts Amherst, rseifried@umass.edu*

Follow this and additional works at: [https://scholarworks.umass.edu/librarian\\_pubs](https://scholarworks.umass.edu/librarian_pubs)

---

### Custom Recommended Citation

Jerome, E., Atwood, T., Radik, M., and R. Seifried. (2023). Envisioning the Future of a Mature IR: A Midlife Assessment of ScholarWorks@UMassAmherst. In Josh Cromwell (Ed.), *Rethinking Institutional Repositories: Innovations in Management, Collections, and Inclusion* (pp. 3-15). Association of College and Research Libraries.

This Article is brought to you for free and open access by the University Libraries at ScholarWorks@UMass Amherst. It has been accepted for inclusion in University Libraries Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

## CHAPTER 1

# Envisioning the Future of a Mature IR:

## A Midlife Assessment of ScholarWorks@ UMassAmherst

*Erin Jerome, Thea Atwood, Melanie Radik, and  
Rebecca Seifried*

### INTRODUCTION

The University of Massachusetts Amherst Libraries launched its institutional repository (IR), ScholarWorks@UMassAmherst, in July 2006. To date, the IR has over 57,000 works that have been downloaded over 25 million times all over the world. Over the past six years, the content of the IR has expanded from mainly postprints and Electronic Theses and Dissertations (ETDs) to include podcasts, datasets, open educational resources (OER), and other shareable open content that has no other logical home. As continued growth has pushed the limits of the software as designed, we decided to conduct a full assessment of the IR in order to qualitatively assess whether the IR meets the needs of staff and stakeholders. The assessment involved two parts: (1) evaluating the IR according to a set of defined criteria adapted from the University of Pennsylvania's Platform



Functionality Review, and (2) soliciting feedback from stakeholders, including those who edit journals or organize conferences (a functionality that is separate from other parts of the IR) and those who manage non-journal or conference collections within the IR. While stakeholders expressed satisfaction with the service provided by the Libraries and the IR provider, they also requested additional features that would make the platform more user-friendly, flexible, and responsive to new content types and customizations that extend beyond the hard limits of the software. Editors and collections administrators were particularly vocal in requesting additional features, such as the ability to accommodate languages other than English and support for big datasets. The assessment raises concerns about the current IR's ability to adapt to changes in the scholarly publishing landscape that are on the horizon and provides critical data to inform the next iteration of the repository. Hopefully, others can apply this strategy to their own institutional repository in order to better prepare for a flexible, robust future that supports open scholarship.

## BACKGROUND

The University of Massachusetts Amherst Libraries launched its institutional repository (IR), ScholarWorks@UMassAmherst, in July 2006 using the commercial hosted IR platform Digital Commons, licensed by bepress. To date, the IR has over 57,000 works that have been downloaded over 25 million times all over the world.<sup>1</sup> While growing steadily since its launch, the types of content collected and showcased in the IR are pushing beyond the traditional IR mainstays of Green Open Access postprints and Electronic Theses and Dissertations (ETDs) to include podcasts, datasets, open educational resources (OER), and other shareable open content. Information about the services provided through ScholarWorks spreads primarily through word of mouth on campus rather than through concerted outreach efforts.

After receiving an increased number of requests to deposit datasets throughout 2016–2017, a data repository was developed and launched within the IR in October 2017.<sup>2</sup> That same year, the IR's publishing capabilities were expanded in response to a request from researchers in the Linguistics department who were exploring possible platforms for hosting an open-access *festschrift*. That publication ultimately resulted in a second *festschrift* for another colleague and laid the groundwork for Linguistics to launch four additional journals/conference proceedings.<sup>3</sup> Once bepress added streaming media support to Digital Commons in 2019, the institutional repository librarian experimented with a faculty member whose final class assignment was to produce a podcast; since then, one additional podcast has been published.<sup>4</sup> The success of these experiments has encouraged more members of the UMass campus to approach the institutional repository librarian with their own unique content.

We have always been experimenters when it comes to the repository, often pushing for customizations and modifications when the platform does not have the needed functionality. While these changes often satisfy short-term needs, there is the risk that future software updates or new feature releases might not support these customizations, thus requiring additional labor to test and debug the software and possibly revert requested changes. After fifteen years of continued growth that has pushed (and at times exceeded) the limits of Digital Commons as a hosted service, IR staff decided to conduct a full assessment of the IR. The goal was to qualitatively assess whether the IR is meeting the needs of staff and stakeholders. The institutional repository librarian and the data services librarian drafted a charge for an IR Assessment Task Force that was approved by the Libraries' leadership in January 2020. With the task force in place, the next step was to develop a plan for assessing the middle-aged IR.

## ASSESSMENT

The assessment began by evaluating the platform's performance with the Platform Functionality Review tool developed by the University of Pennsylvania's Scholarly Communication and Research Infrastructure Project (SCRIP).<sup>5</sup> This publicly available tool (CC BY 4.0) includes details of several major platforms, making it possible to assess how the IR performed in the broader context of available services. It offers short descriptions of functionality for several areas of both IR and journal hosting, providing a granular and flexible assessment of usability for content managers and users. The sixty questions in the IR assessment section and forty-eight questions in the journal platform section were narrowed down to a dozen must-have functions that allowed the team to determine whether Digital Commons was meeting our institutional needs.<sup>6</sup>

The second phase of the assessment involved soliciting stakeholder feedback. Based on insights from the platform review, support tickets, and user comments during consultations, a set of targeted interview questions were created to add users' perspectives on the platform's performance (see appendix A). Since this phase took place at the height of the COVID-19 pandemic, interviews were performed asynchronously by email rather than via Zoom and interview questions were limited to the most salient topics. The team reached out to power users, including those who edit journals or organize conferences and those who manage non-journal or conference collections within the IR. In total, there were ten responses to the questionnaire: five from journal editors or conference organizers and five from other stakeholders.

## RESULTS

The assessment provided insights into the community's successes and frustrations with the IR platform. Stakeholders expressed satisfaction with the service as the service

provider as well as with the people “behind the scenes” who manage the resources, metadata, and software customizations of the IR—the institutional repository librarian and the data services librarian. They also requested additional features to make the platform more user-friendly, flexible, and responsive to new content types.

## *Needs the Platform Is Meeting*

While users regularly deposit content that is traditionally associated with IRs (e.g., text-based items like journal articles, theses and dissertations, working papers, etc.), non-text-based collections have begun to grow in the past few years. Digital Commons natively supports text-based items, and the addition of streaming media functionality has enabled experimentation with audio and video files. We have also implemented customizations or used workarounds to allow users to deposit other types of resources, like podcasts and datasets—the latter of which requires using Digital Commons’ predefined “book” format to display properly in the user interface.

Stakeholders frequently cited digital object identifiers (DOIs) as a reason why they use ScholarWorks instead of resources such as academia.edu or a personal website. The Libraries rely on a DataCite membership to mint DOIs for much of the content that is uploaded to the IR (e.g., ETDs, datasets, and library-published journal articles), using the DataCite application programming interface (API) or Fabrica interface to mint a DOI, and then manually adding it to an item’s metadata record. Groups of items (like ETDs and journals) can be edited with Digital Commons’ batch revise tool, which generates an Excel spreadsheet containing the metadata of all items in a specific publication structure that administrators can then edit and upload back into Digital Commons.<sup>7</sup> Thus, while DOIs are a valuable aspect of the IR, this service does not rely on Digital Commons.

In addition to these custom workflows, the assessment identified four strengths of the core Digital Commons platform: (1) journal publishing capability, (2) a robust analytics dashboard, (3) ability to enable embargoes and access restrictions for ETDs, and (4) search engine optimization (SEO). One of the benefits of Digital Commons is that it functions as both a repository and a journal publication platform, effectively serving as an all-in-one platform. This has enabled the library publishing program to develop and grow, including migrating journals from other platforms to ScholarWorks. Journal editors appreciate the editorial features associated with the journal publishing suite of tools, such as the support for submission, double-blind peer review, and publication all in a single platform.

The analytics dashboard featured prominently in stakeholder feedback (figure 1.1 and figure 1.2). It includes monthly readership reports and real-time download statistics for all items associated with a journal and/or user account. Both stakeholder communities commented on the usefulness of the analytics, and journal editors specifically mentioned that their authors liked this feature. In



decision in 2014 to mandate deposits to the IR instead of ProQuest. In particular, the Graduate School required the ability to apply campus-only access restrictions that are functional for both residential and off-campus UMass users.

Users often ask if their work will be findable if they upload it to ScholarWorks. The IR is indexed by all the major search engines, thanks to Digital Commons' out-of-the-box search engine optimization (SEO) features that are meant to optimize content for Google, Google Scholar, and other popular search engines.<sup>9</sup>

## User-Identified Gaps in Functionality

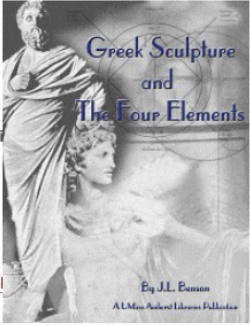
As a hosted solution, Digital Commons is understandably unable to accommodate all possible customization requests since every software customization has the potential to complicate future collections or may not be supported in future software upgrades (figure 1.3 and figure 1.4). Nevertheless, stakeholders expressed

### Greek Sculpture and the Four Elements

**Greek Sculpture and the Four Elements [full text, not including figures]**

**Contents**

- ▣ » [Greek Sculpture and the Four Elements \[full text, not including figures\]](#)
- ▣ [Cover and front matter](#)
- ▣ [Preface](#)
- ▣ [Introduction](#)
- ▣ [Chapter 1: Toward defining the ego](#)
- ▣ [Chapter 2: Survey of examples of period-setting in studies of Greek \(or Greek sculpture\) in modern scholarship](#)
- ▣ [Chapter 3: Human consciousness and time](#)
- ▣ [Chapter 4: Form and time](#)
- ▣ [Chapter 5: A detailed study of the emergence and significance of contrapposto](#)
- ▣ [Epilogue: On philosophers and artists](#)
- ▣ [Appendix A: Diagrams illustrating the progressive governance of the four elements-cycle by each element in turn](#)
- ▣ [Appendix B: The structure of a world view \(Kluback & Weinbaum 1957, 25-27\)](#)
- ▣ [Appendix C: Triadism in the history of 19th and early 20th century psychology](#)
- ▣ [Appendix D: On the question of mineral weight in the Canon](#)
- ▣ [Select bibliography](#)
- ▣ [Notes](#)
- ▣ [Figures](#)



**Download**

**SHARE**













[f](#)
[t](#)
[in](#)
[e](#)
[+](#)

**Figure 1.3**

Screenshot of a now-unsupported customization that mimicked an e-book table of contents.

the desire to customize or modify almost every aspect of its interface, from the overall appearance of landing pages to the ability to edit submission forms and decision letters that are automatically sent as part of the journal review process. Editors working with multilingual content and non-English-speaking authors have also expressed frustration that many aspects of the Digital Commons back-end, submission, and peer-review tools are hard-coded in English.

#### Digital Books from 2000

-  PDF [Appendix A: Diagrams illustrating the progressive governance of the four elements-cycle by each element in turn](#), J.L. Benson
-  PDF [Appendix B: The structure of a world view \(Kluback & Weinbaum 1957, 25-27\)](#), J.L. Benson
-  PDF [Appendix C: Triadism in the history of 19th and early 20th century psychology](#), J.L. Benson
-  PDF [Appendix D: On the question of mineral weight in the Canon](#), J.L. Benson
-  PDF [Chapter 1: Toward defining the ego](#), J.L. Benson
-  PDF [Chapter 2: Survey of examples of period-setting in studies of Greek art \(or Greek sculpture\) in modern scholarship](#), J.L. Benson
-  PDF [Chapter 3: Human consciousness and time](#), J.L. Benson
-  PDF [Chapter 4: Form and time](#), J.L. Benson
-  PDF [Chapter 5: A detailed study of the emergence and significance of contrapposto](#), J.L. Benson
-  PDF [Cover and front matter](#), J.L. Benson
-  PDF [Epilogue: On philosophers and artists](#), J.L. Benson
-  PDF [Figures](#), J.L. Benson
-  PDF [Greek Sculpture and the Four Elements \[full text, not including figures\]](#), J.L. Benson
-  PDF [Introduction](#), J.L. Benson
-  PDF [Notes](#), J.L. Benson
-  PDF [Preface](#), J.L. Benson

#### Figure 1.4

Screenshot of the same structure with customization removed. Contents are now ordered alphabetically by title.

The assessment revealed that users want a good deal of control over how their work is presented, and they want to be able to make changes without having to ask for assistance from the Libraries or bepress. Some users noted that even when they can customize part of an item's presentation, these customizations may not function as expected. For example, authors can create a custom citation, but it



will not appear on the item's landing page without asking bepress to suppress the automatically generated citation. By comparison, non-institutional repositories like Zenodo allow users to configure the recommended citation instantly—a functionality that one stakeholder mentioned as a reason for continuing to submit their research products there instead of with ScholarWorks. One potential strength of the Digital Commons platform is the integration between ScholarWorks and the Expert Gallery Suite (EGS, a platform that allows authors to curate their research content on a personalized webpage); ideally, EGS provides the kind of individual control that scholars want. Unfortunately, stakeholders reported confusion about how to log in to the platform due to ambiguous labeling in the site's navigation and differing sign-in requirements. Some authors even mistake EGS for ScholarWorks and upload content to their profile instead of depositing it in the IR.<sup>10</sup>

A source of frustration for both users and administrators is ScholarWorks' lack of integration with ORCID, Crossref, DataCite (the Libraries' DOI provider), and big data infrastructure like Globus. bepress has developed two versions of an API that administrators can use to analyze metadata, create annual reports, and (with more programming skills) integrate with other services. However, unlike other IR platforms, Digital Commons does not have built-in plugins for easily retrieving or sharing data such as DOIs and ORCIDs.

Users also indicated a desire to populate ScholarWorks collections with information from their curricula vitae or annual reports. Such functionality would provide seamless support for researchers and improve the campus's adoption of the IR. While bepress has released metadata-harvesting tools that work with ORCID, PubMed, and Scopus, the greatest need is for pipelines that share metadata and downloadable content between systems. Despite its age, Simple Web-service Offering Repository Deposit (SWORD) is considered a well-established deposit protocol, yet it remains incompatible with Digital Commons.

Another aspect of Digital Commons that creates additional labor for administrators is the inability to manage user accounts locally; only bepress can manage user accounts or troubleshoot when an account issue arises. ScholarWorks is set up so that any user can create an account, a functionality considered necessary because the IR's journal and conference proceedings allow submissions from beyond campus and because graduate students' email addresses do not persist after graduation. However, IR staff are unable to impersonate user accounts or view the email addresses associated with user logons. Single sign-on (SSO) is supported, but not Shibboleth, which the campus currently requires. This has been a continued source of frustration for users and administrators. For example, the inability to manage, merge, or de-duplicate accounts can make for a great deal of redundant work and confusion for graduate students, who are required to submit their dissertations and theses to ScholarWorks.

The final gap in functionality is related to the age of the IR and specifically to the fact that many different individuals—from both UMass and bepress—have managed ScholarWorks over the years. The consequences of these changes in oversight and administration are especially apparent in metadata. The benefits of well-described metadata that follow established standards include improved findability, reuse, and long-term preservation—benefits that are often invisible to end users but that nevertheless enhance their experience of using the IR. Unfortunately, while Digital Commons’ metadata can be mapped to both Dublin Core and a custom export label, these mappings have never been standardized in any consistent way in ScholarWorks. Most recently, an attempt to reintegrate ETD metadata via OAI harvesting into the Libraries’ discovery layer (EBSCO Discovery Service) revealed a mistaken mapping of the Dublin Core document-type field, which required clean-up and remapping within the ETD collections. Metadata mapping and inconsistencies also likely play a role in ScholarWorks’ surprising inability to search itself. In short, the stakeholder feedback made clear that changes in IR administration, as well as a lack of accessible UMass-specific documentation and policies since its inception, have negatively impacted users’ experiences.

## NEXT STEPS

Two major themes emerged during the assessment: first, the tension between adding customizations to meet users’ needs and “future-proofing” for anticipated software upgrades, and second, the fact that there is no perfect platform that can support every possible need users might have. Nevertheless, the assessment highlighted both areas where the current IR is doing well and areas where it could be improved. Hopefully, these findings will contribute to library- and community-led development efforts toward new open-source solutions that suit UMass’s continuously evolving needs.

Importantly, in addition to the strengths and weaknesses of the chosen IR platform, the overall repository infrastructure is also subject to friction created by the human component. ScholarWorks was created in 2006, but best practices for IRs did not surface until years later.<sup>11</sup> While being on the cutting edge gave the institution the freedom to explore, it also resulted in a degree of chaos that newer IRs may not experience. This IR also reflects fifteen years’ worth of working with bepress consultants, who brought their own unique experience and understood to varying degrees the need to fully discuss the possible impacts of changes made to the repository. As a result, ScholarWorks itself is an amalgamation of priorities, desires, and interpretations, and a great deal of work is required to corral the excitement of the many hands that have touched it.

So, where to go from here? It is necessary to develop workflows and policies that guide approaches to modifications, always with an eye toward their possible

impact on future functionality. For example, customizations are difficult to maintain over the long term, but no product will meet all users' needs out of the box. Perhaps when users approach us with content that cannot be accommodated without a great deal of customization, the difficult decision must be made to refer that content owner to another platform.

Armed with the knowledge that no single platform can possibly do everything, it is important to look for a platform that checks as many boxes on the list of desired functionality and features as possible. Even within that list, the critical features identified in the assessment must be prioritized: the ability to publish journals (and give editors the ability to customize forms related to the review process); provide download statistics and analytics data; enable auto-lifting embargoes and access restrictions for ETDs; optimize search-engine results; allow users to customize the way their content appears; integrate with platforms like ORCID and DataCite; manage user accounts; and standardize metadata mapping with well-established schemas.

This assessment raises concerns about the current IR's ability to adapt to changes in the scholarly publishing landscape that are on the horizon, and it provides critical data to inform the next iteration of the repository. Hopefully, others can apply this strategy to their own institutional repositories in order to better prepare for a flexible, robust future that supports open scholarship.

# Appendix A. Stakeholder Feedback Questions

For non-journal/conference stakeholders:

1. What have you used ScholarWorks for?
2. What do you like about the ScholarWorks platform (features, creating collections, review workflows, statistics dashboard, etc.)?
3. Have you ever been disappointed by the ScholarWorks platform? If so, why?
4. Are there aspects of ScholarWorks' functionality that have caused you to find other platforms for specific kinds of content? If so, what other platforms?
5. What kind of materials do you share right now? Do you anticipate sharing other types of materials in the future?
6. What kind of functionality would you want to see in an ideal platform?

For journal editors/conference organizers:

1. How would you describe your experience setting up a new journal or conference?
2. Are there parts of the creation process that you would like to change?
3. What do you like about the ScholarWorks platform (peer review, editorial workflows, statistics, submission process)?
4. Are there pain points or features of ScholarWorks that you wish you could change? If so, why?
5. What kind of functionality would you want to see in an ideal platform?

## NOTES

1. ScholarWorks dashboard, 07/19/2006–08/24/2023.
2. Thea P. Atwood, Erin Jerome, Ann Kardos, Stephen McGinty, Melanie Radik, and Rebecca Reznik-Zellen, "Cross-Functional Policy Development for the UMass Amherst Data Repository," *Journal of Librarianship and Scholarly Communication* 9, no. 1 (2021), <https://doi.org/10.31274/jlsc.12911>.
3. Nicholas LaCara, Keir Moulton, and Anne-Michelle Tessier, "A Schrifft to Fest Kyle Johnson," *Scholarworks@UMassAmherst: Linguistics Open Access Publications* 1 (2017), <https://doi.org/10.7275/R57D2S95>; Rajesh Bhatt, Ilaria Frana, and Paula Menéndez-Benito, "Making Worlds Accessible. Essays in Honor of Angelika Kratzer," *ScholarWorks@UMass Amherst: Angelika Kratzer Festschrift* 1 (2020), <https://doi.org/10.7275/w10a-pv24>.
4. Published podcasts: <https://scholarworks.umass.edu/finalexamination/>; <https://scholarworks.umass.edu/trajectory>.
5. "Platform Functionality Review," Scholarly Communication and Research Infrastructure Project (SCRIP), University of Pennsylvania, posted January 23, 2019, <https://penntnl.wordpress.com/2019/01/23/scríp-appendices/>.

6. Erin Jerome, Thea Atwood, Melanie Radik, and Rebecca Seifried, "Assessment Data for "Envisioning the Future of a Mature IR," *Scholarworks@UMassAmherst: Data and Datasets* 149 (2022), <https://doi.org/10.7275/hsrc-4b43>.
7. "Batch Upload, Export, and Revise," bepress, accessed 4 October 2022, [https://bepress.com/reference\\_guide\\_dc/batch-upload-export-revise/](https://bepress.com/reference_guide_dc/batch-upload-export-revise/).
8. We have two moving-wall access restrictions in place: 1-year and 5-year. However, the software can only accommodate one moving wall per publication structure. Our workaround is to set the 5-year restriction as the default for automatic lifting, and then if a user opts for the 1-year restriction, administrators in the Libraries must manually lift the embargo after one year.
9. "Search Engine Optimization: Features and Best Practices," bepress, accessed April 28, 2022, [https://bepress.com/reference\\_guide\\_dc/search-engine-optimization-repository/](https://bepress.com/reference_guide_dc/search-engine-optimization-repository/).
10. Much of this confusion is due to the fact that EGS used to be called SelectedWorks, a name very similar to ScholarWorks.
11. See, for example, Pamela Bluh and Cindy Hepfer, *The Institutional Repository: Benefits and Challenges* (Chicago: American Library Association, 2013); Burton B. Callicott, David Scherer, and Andrew Wesolek, *Making Institutional Repositories Work* (West Lafayette, IN: Purdue University Press, 2015); Jonathan A. Nabe, *Starting, Strengthening, and Managing Institutional Repositories: A How-To-Do-It Manual* (New York: Neal-Schuman Publishers, 2010).

## BIBLIOGRAPHY

- Atwood, Thea P., Erin Jerome, Ann Kardos, Stephen McGinty, Melanie Radik, and Rebecca Reznik-Zellen. "Cross-Functional Policy Development for the UMass Amherst Data Repository." *Journal of Librarianship and Scholarly Communication* 9, no. 1 (2021): eP12911. <https://doi.org/10.31274/jlsc.12911>.
- Bhatt, Rajesh, Ilaria Frana, and Paula Menéndez-Benito. "Making Worlds Accessible. Essays in Honor of Angelika Kratzer." *ScholarWorks@UMass Amherst: Angelika Kratzer Festschrift* 1 (2020). <https://doi.org/10.7275/w10a-pv24>.
- Bluh, Pamela, and Cindy Hepfer. *The Institutional Repository: Benefits and Challenges*. Chicago: American Library Association, 2013.
- Callicott, Burton B., David Scherer, and Andrew Wesolek. *Making Institutional Repositories Work*. West Lafayette, IN: Purdue University Press, 2015.
- Jerome, Erin, Thea Atwood, Melanie Radik, and Rebecca Seifried. "Assessment Data for "Envisioning the Future of a Mature IR." *Scholarworks@UMassAmherst: Data and Datasets* 149 (2022). <https://doi.org/10.7275/hsrc-4b43>.
- LaCara, Nicholas, Keir Moulton, and Anne-Michelle Tessier. "A Schrif to Fest Kyle Johnson." *Scholarworks@UMassAmherst: Linguistics Open Access Publications* 1 (2017). <https://doi.org/10.7275/R57D2S95>.
- Nabe, Jonathan A. *Starting, Strengthening, and Managing Institutional Repositories: A How-To-Do-It Manual*. New York: Neal-Schuman Publishers, 2010.
- University of Pennsylvania. "Platform Functionality Review." *Scholarly Communication and Research Infrastructure Project (SCRIP)*. Posted January 23, 2019. <https://penntnl.wordpress.com/2019/01/23/scrif-appendices/>.