Developing a Personalized Institutional Repository Framework: How the Colorado School of Mines Adapted Repository Certification Criteria around Local Needs Developing a Personalized Institutional Repository Framework to Address Local Priorities and Goals

Emily Bongiovanni
emilybongiovanni@mines.edu

Torey Battelle
battelle@mines.edu

Beth Zecca
ezecca@mines.edu

Colorado School of Mines
Colorado, USA

Abstract

Purpose
The purpose of this paper is to document the challenges and lessons learned while developing a local international repository (IR) framework at the Colorado School of Mines (Mines). The aim of this exercise was to characterize IR needs, to understand how successfully these needs are being met, and to develop goals for future IR development. The framework was developed with the intention of being used as an instrument for periodical self-assessments, to observe the IR development, and to aid in strategic decision-making.

Method
The Mines IR Steering Committee used criteria and guidelines from digital repository certificates, including the Trusted Repositories Audit Certification and CoreTrustSeal, and existing frameworks to develop a local framework tailored to Mines unique institutional needs.

Findings
This exercise demonstrated the range of perspectives and priorities for the IR among the IR Steering Committee. A brief assessment that tested the framework highlighted the current condition of the IR at Mines and areas for improvement.

Practical Applications
Librarians and professionals working with IRs can apply observations and lessons learned from the framework development process.

Statement of originality
This paper offers a case study for developing a local IR framework to help self-assess, strategize, and better characterize ongoing needs for the development of a high-quality IR. Librarians and professionals working with IRs can apply lessons learned to help create a local IR framework at their institution.

Keywords: Institutional repository, IR development, Assessment

Introduction

The Colorado School of Mines (Mines) institutional repository (IR) serves as a stable, digital platform for Mines students, faculty, and groups to disseminate their research, projects, and other works. Mines benefits from a shared consortial institutional repository, called Mountain Scholar. Mines' participation with Mountain Scholar is managed by the IR Steering Committee, which includes library faculty and staff and a member of the Research Support Services group.

Through conversation, the Mines IR Steering Committee saw a need to characterize goals for the IR. The Committee recognizes that Mountain Scholar is providing Mines with a low-cost and effective IR solution. However, to continue meeting the Mines research community's needs, characterization and prioritization of IR needs and goals is necessary. The IR Steering Committee attempted to assess how Mountain Scholar is addressing Mines' IR needs by developing a framework inspired from the criteria and guidelines from official repository certificates and previous developed frameworks. The framework that the IR Steering Committee developed, called the Framework for the Institutional Repository at Mines (FIRM), is intended to help the Committee self-assess, strategize, and better characterize ongoing needs for the development of a high-quality IR. This paper discusses the challenges and lessons learned during the process of creating the FIRM. Other institutions can apply lessons learned to the development and maintenance of their own IR.

Review of Literature

An institutional repository (IR) is either defined as a platform to manage and share an institution's output or as the services related to disseminating an institution's work in their digital repository. Johnson (2002) explains an IR as "a digital archive of the intellectual product created by the faculty, research staff, and students of an institution and accessible to end users both within and outside of the institution" (para 12). However, Lynch (2003) defines an IR as "a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members" (pg. 328). Lynch's definition demonstrates the collaborative effort from different institutional units in the stewardship, preservation, organization, access, and distribution of content. Both definitions are relevant to the IR in the Mines context. Gibbons (2004) outlines the core functions of an IR: material submission; metadata application; access control; discovery support; distribution; preservation. An IR is valuable to its research community because it can significantly improve access for a wider and more diverse audience and enable long-term preservation (Swan, 2008).

The success of an IR can be significant to an institution's scholarly footprint: Johnson (2002) explains, "institutional repositories, by capturing, preserving, and disseminating a university's collective intellectual capital, serve as meaningful indicators of an institution's academic quality" (para. 8). The group managing the IR must be able to identify the factors needed to provide a successful IR.

Lagzian, Abrizah, and Wee (2015a) explain "there has been no comprehensive study on the success factors of IRs in the library and information science context, as such, no investigation of the factors in some real-life contexts of IR activities" (pg. 197). However, there are a few frameworks on

conditions for a successful IR. For example, Westell (2006) considers eight indicators of IR success: mandate, integration with planning, funding model, relationship with digitization centers, interoperability, measurement, promotion, and preservation strategy. Thibodeau (2007) proposes success factors comprising of service, chronological orientation, coverage, collaboration, and state of development. Swan (2008) suggests a framework for performance indicators that consists of content recruitment, user awareness and involvement, workflow practices, and financial discipline. Literature on conditions that make an IR successful emphasize content recruitment and user participation (Lagzian, et al., 2015a). In fact, Lagzian, Abrizah, and Wee (2015b) stress self-archiving practices as one of the most important "critical success factors." In addition, case studies emphasize the importance of assessing the research community's needs in order to best align IR policies, procedures, and strategies (Cassella, 2010).

There are certifications that digital repositories can obtain to indicate their quality, such as The Trusted Repositories Audit and Certification: Criteria and Checklist (TRAC) and CoreTrustSeal. These certificates are not designed specifically for institutional repositories. However, they address fundamental components of high-quality digital repositories that can be applied to the IR context. TRAC was developed by the Center for Research Libraries as a tool to measure a digital repository's ability to "preserve digital content in a way that serves the repository's stakeholder community" (Center for Research Libraries, n.d., para. 1). The Audit and Certification of Trustworthy Digital Repositories: Recommended Practice is a document which provides recommended practices and standards based off TRAC (The Consultative Committee for Space Data Systems, 2011). It includes criteria regarding organizational infrastructure, digital object management, and infrastructure and security risk management.

The Core Trustworthy Data Repository Certificate, also called the CoreTrustSeal (previously Data Seal of Approval), is a digital repository certification that aims at sustainable and effective data sharing practices and infrastructure (CoreTrustSeal, 2017a). The CoreTrustSeal certification has 16 mandatory requirements that fall under three broader categories: Organizational Infrastructure, Digital Object Management and Technology.

The CoreTrustSeal is the result of a merger of the World Data System of the International Science Council and the Data Seal of Approval (CoreTrustSeal, 2017a). There are currently 53 digital repositories around the world that are certified with this new CoreTrustSeal (CoreTrustSeal, 2017b). The majority of CoreTrustSeal certified repositories are discipline-based repositories, such as the United States' National Snow & Ice Data Center. There are a few CoreTrustSeal certified IRs, including the Missouri University of Science and Technology's Scholars Mines and University of California's Merritt.

The CoreTrustSeal is especially relevant to data sharing mandates. Many of the requirements of these certificates align with data repository requirements from grant-funding agencies. Major grant funding agencies in The United States, including the National Science Foundation (NSF), the Department of Energy (DOE), and the National Institutes of Health (NIH) have made significant changes recently to now require grantees to make their research data openly available. In addition to agency-wide commitments and mandates to make publicly funded research open, some individual grants have supplementary requirements. For example, NSF grant 19-528 states '[the data] must be in a national, public data repository that publishes its contents to a higher-level data

aggregator that facilitates data discovery. This public repository should be registered with an international body that promotes best practices in data archiving and curation' (pg. 9). Researchers at Mines are funded by grants such as NSF grant 19-528 which specifically outline requirements to curate, share, and archive research data.

IR at Mines

Colorado School of Mines (Mines) is a medium-sized public university focused on disciplines related to science, technology, engineering and mathematics. Mines participates in a consortia-based IR called Mountain Scholar. It is shared with other academic institutions in Colorado and Wyoming and managed by Colorado State University, Fort Collins. Mountain Scholar is hosted on a DSpace platform, which is a free, open source software that academic, other non-profit, and commercial organizations use to build digital repositories (Duraspace, n.d.).

The Arthur Lakes Library has a small team of librarians and there is no one individual whose job is primarily focused on the IR. Instead, the Scholarly Communications Librarian chairs the IR Steering Committee, which is made up of various library faculty and staff and a member of the university's Research Support Services group. The IR Steering Committee manages the policies, procedures, and workflows regarding Mines' participation in Mountain Scholar.

The primary mission of the IR is to make Mines digital content openly available online and to preserve it for future access and use. Mines' IR content consists of works created by or in collaboration with Mines faculty, staff, and students, as well as works from Mines' Special Collections and Archives. It includes works that were born digital and items that have been digitized, and includes text, images, papers, data sets, videos, to maps. Content falls under the following categories:

- Research and scholarly outputs
- Educational and informational works
- Unique works of the university and its departments
- Creative and recreational works generated as part of Mines programs

Mines' current IR content consists mostly of theses, dissertations, and items from the Special Collections. There are various other items, however faculty research content in the IR is relatively small part of the content.

Research at Mines centers around earth, energy and the environment, and covers a variety of topics related to science, technology, and engineering. Mines is awarded a large amount of financial research support for its size from funding agencies. In fiscal year 2018, the institution received $65.89 million in research awards. Many of these funds come from agencies with mandates to share research data, such as NSF, NIH, and DOE. A high-quality IR could support many of the research community's mandates to share research data.

Purpose

The IR Steering Committee recognizes that Mines' IR does not currently meet all of Mines' research community's needs. The Committee therefore found it necessary to collectively outline the fundamental components of a successful IR for Mines.

The IR Steering Committee adapted official repository certification criteria and existing IR frameworks to develop a personalized framework for Mines. The purpose of this exercise was to characterize IR needs, to understand how successfully these needs are being met, and to develop goals for future IR development. The framework was developed with the intention of being used as an instrument for periodical self-evaluations to observe the IR development and as a document to aid in strategic decision making.

Method

The IR Steering Committee formed a working group to lead the development of the Framework for the Institutional Repository at Mines (FIRM). This group included the Scholarly Communications Librarian, the Assistant Director of Research Computing, and the library's Digital Initiatives Specialist. This combination of team members brought various perspectives and priorities on research dissemination and information management.

The working group collaboratively worked to develop the FIRM's two sections: a section of necessary, fundamental components of the IR and a section of features that would make the IR unique and exceedingly successful. The group considered the IR as both a platform and a set of services when developing the FIRM. No ranking or prioritization was given to items besides placement in the two categories, however the group anticipates establishing this in future versions of the FIRM.

The working group first reviewed the recommended practices and standards provided by the Audit and Certification of Trustworthy Digital Repositories and the 16 requirements for the CoreTrustSeal. Each member of the working group went through the two documents and marked criteria in either of the two categories or to be omitted as unnecessary for Mines. In addition, the group looked at other IRs that are certified by either the CoreTrustSeal or Audit and Certification of Trustworthy Digital Repositories to observe how these requirements are applied and to discover other characteristics that could be valuable for the Mines IR. The group also reviewed existing literature on IR development and success factors and a selection of Mines' pertinent funding agencies' sharing mandates.

The FIRM working group brought the document to the IR Steering Committee for discussion and feedback. This ensured each member of the Committee's perspective are reflected in the FIRM.

Findings

The FIRM identifies critical components the IR needs to serve its purpose at Mines. As Thibodeau (2007) proposes, a framework must reflect the IR's purpose because "no repository can be said to be truly successful in a meaningful sense unless it fulfills its purpose. Thus, criteria for success must be derived from its statement of purpose" (para. 2).
The FIRM content reflects most of the criteria outlined in the CoreTrustSeal and the Audit and Certification of Trustworthy Digital Repositories requirements, as well as the existing IR frameworks. The FIRM covers various IR aspects, including internal workflows, features for discoverability, and preservation policies.

Items in the FIRM are placed into two categories: *core components* that the Committee finds necessary to have for a proper, manageable, functioning digital repository; and *aspirational features* that the Committee believes will enable the IR to be a unique and advantageous resource for the Mines community. The *core components* for the IR comprise the following:

- Mission statement which reflects our commitment to preservation and access
- Collection policy
- Continuity plans
- Strategic preservation plan
- Operating workflow
- Digital infrastructure
- Rich metadata
- Security protection and backup up processes

This *core components* section reflects core aspects for an adequately functioning IR. This section aligns with Mines' IR commitment to share Mines' work with world (mission statement, collection policy, rich metadata, digital infrastructure) and to preserve content over time (continuity plans, strategic preservation plan, security protection and back up processes). This section also addresses aspects regarding long-term preservation, security risks, and infrastructure changes.

The *aspirational features* section outlines elements that would make the IR unique among academic institutions and of particularly high quality. Items in this category include:

- Self-submission capabilities (self-archiving)
- Features to enhance scholarly footprint
- Processes that promote best practices in sharing research data
- Discovery and identification features
- Restricted access controls
- Outreach and marketing support

This section encompasses features such as automatic citations and bi-directional linage between related items or resources. The *aspirational features* list unintentionally reflects features that serve end-user's interaction with the IR.

The FIRM document details each item in greater extent by providing examples of features.

Discussion

The working group members' three different lenses provided differences in perspective of the framework. For example, the Assistant Director of Research Computing, who looks at research data management plans daily, put emphasis on the criteria related to data quality and review. Whereas, the Scholarly Communications Librarian put emphasis on criteria related to discoverability and impact while the Digital Initiatives Specialist focused on content types and formats. These varying perspectives led to rich discussions and a comprehensive FIRM that would not have happened had only one member of the Committee worked on the FIRM.

One of the greatest challenges was determining between the two categories for each item. For example, the group was challenged in determining if self-submission capabilities should be characterized under *core components* or *aspirational features.* While most of literature on IR frameworks emphasize self-submissions, the group did not consider it critical to the foundation of the IR. This decision was influenced by the current status of self-submission capabilities at Mines: The Mines IR does not currently provide a self-submission feature, however the group finds the IR is functioning adequately without it.

Most of the requirements from the CoreTrustSeal and the Audit and Certification of Trustworthy Digital Repositories are included in the FIRM in adapted or identical forms. Requirements from the certificates that were omitted from the FIRM include criteria regarding managing licenses, managing contracts, staffing, and ensuring data shared in the IR was created/curated in compliance with disciplinary/ethical norms. These omitted items were excluded because they do not fall under the IR's current priorities or mission. For example, the Committee omitted processes to ensure research data shared in the IR was created according to disciplinary and ethical norms (Requirement 4 of the CoreTrustSeal) because the IR's current collection policy does not limit content based on quality or disciplinary/ethical norms. The group did not consider changing the collection policy to reflect these content controls because they believe this could potentially prevent early career researchers who are still learning their discipline's best practices from sharing their works on the IR.

There are some items in the FIRM that are not included in the Audit and Certification of Trustworthy Digital Repositories or CoreTrustSeal. These items were included because they were identified during the review of literature or are already known to be particularly important to the Mines context. For example, the Scholarly Communications Librarian recognized the ability to utilize the IR to enhance scholarly footprint. Features related to scholarly impact, such as providing citable links and integrating with systems that track research impact, could support the reuse and visibility of works. As Jean, Rieh, Yakel, and Markey (2011) explain, the ability to demonstrate impact is critical for the recruitment of participation from content submitters. The Committee believes that features to help increase citations and visibility and also follow usability and discoverability best practices can increase IR participation.

The most significant lessons learned from this exercise are 1) the range of perspectives among the IR Steering Committee, and 2) the need for community feedback. The working group did not expect such varying priorities among the members and were grateful for the rich discussions that resulted from the various perspectives. The working group also recognized that the MIRF cannot be successful without feedback from users. Each of the working group's discussions included an unanswered question regarding depositor or audience interactions with the IR.

Limitations

There are various limitations to this first attempt at an IR framework. The most significant limitation is the lack of user feedback. As highlighted by Cassella (2010), pre-implementation qualitative assessment is essential in understanding what an institution's research community needs and wants from their IR. The group plans to obtain user feedback from the IR's audience

and clients through surveys and focus groups in the future. This exercise would include investigating the usability of the IR from the audience's perspective, in addition to the depositors' need. The Committee suspects that the incorporation of community feedback will result in more even distribution of back-end and user-end features in the two categories.

The working group also did not examine all of the sharing mandates from the Mines research community' major funding agencies. The Committee will need to take a comprehensive look at a greater variety of funder dissemination mandates and the best practices for sharing research among the community's disciplines to make better informed decisions.

Finally, the group also needs to explore more literature and case studies on IRs to better understand the challenges and opportunities frameworks and certificates.

Application

The IR Steering Committee used the FIRM to complete a brief assessment of the current condition of the IR at Mines. This preliminary assessment demonstrated areas the IR already meets and areas where the IR can improve. This assessment highlighted *core components* that the IR does not currently have in place, such as continuity plans and a documented workflow. Surprisingly, the Committee found that the group's current plans for IR development reflect the FIRM's *aspirational features* more than the *core components.*

The IR Steering Committee will continue to use the FIRM for periodical self-assessments to ensure the group is continuously working on maintaining and advancing the IR according to Mines' needs. The FIRM may also help direct funding priorities and promote additional financial support from the institution. The FIRM will also be used to help facilitate conversations with Colorado State University and other members of the consortium when discussing shared goals and priorities.

This paper and the FIRM can be used by institutions to help develop their own IR framework. Librarians and professionals working with IRs can apply lessons learned to help create a diverse working group and to ensure community feedback is incorporated.

Conclusion and future study

The FIRM is a first attempt at an IR framework for Mines. It can be used by other institutions with similar size, research programs, funding agency mandates, and aspirations in disseminating scholarly output. The IR Steering Committee found it beneficial to create the FIRM as collective effort. This exercise facilitated discussion among the group and highlighted Committee members' varying priorities and strategic ideas for the IR.
There are various other approaches the Committee can take to capture more input on the research community's need to strengthen the FIRM. The Committee intends to improve the FIRM with community feedback before disseminating.

References

Cassella, M. (2010). Institutional repositories: an internal and external perspective on the value of IRs for researchers' communities. *Liber Quarterly*, *20*(2), 210–225. DOI: http://doi.org/10.18352/lq.7989

Center for Research Libraries. (n.d.). TRAC metrics. Retrieved from https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/trac

The Consultative Committee for Space Data Systems. (2011). *Audit and certification of trustworthy digital repositories*. 77.

CoreTrustSeal. (2017a, June 27). About. Retrieved from https://www.coretrustseal.org/about/

CoreTrustSeal. (2017b, June 28). Core certified repositories. Retrieved from CoreTrustSeal website: https://www.coretrustseal.org/why-certification/certified-repositories/

Duraspace. (n.d.). About DSpace - DSpace. Retrieved from https://duraspace.org/dspace/about/

Gibbons, S. (2004). Defining an institutional repository. *Library Technology Reports*, *40*(4), 6–10. Retrieved from https://journals.ala.org/index.php/ltr/article/view/4378/5061

Jean, B. S., Rieh, S. Y., Yakel, E., & Markey, K. (2011). Unheard voices: institutional repository end-users. *College & Research Libraries*, *72*(1), 21–42. DOI: https://doi.org/10.5860/crl-71r1

Johnson, R. (2002). Institutional repositories: Partnering with faculty to enhance scholarly communication. *D-Lib Magazine*, *8*(11). Retrieved from http://www.dlib.org/dlib/november02/johnson/11johnson.html

Lagzian, F., Abrizah, A., & Wee, M. C. (2015a). Critical success factors for institutional repositories implementation. *The Electronic Library*, *33*(2), 196–209. https://doi.org/10.1108/EL-04-2013-0058

Lagzian, F., Abrizah, A., & Wee, M.-C. (2015b). Measuring the gap between perceived importance and actual performance of institutional repositories. *Library & Information Science Research*, *37*(2), 147–155. https://doi.org/10.1016/j.lisr.2014.06.007

Lynch, C. A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy*, *3*(2), 327–336.

National Science Foundation. (2018). *CNH2: Dynamics of Integrated Socio-Environmental Systems (CNH2)*. National Science Foundation.

Swan, Alma (2008). The business of digital repositories. In, Weenink, Kasja, Waaijers, Leoand van Godtsenhoven, Karen (Eds.) A DRIVER's *Guide to European Repositories (Amsterdam, 2007).* Amsterdam University Press.

Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository. *Journal of Digital Information*, *8*(2). Retrieved from https://journals.tdl.org/jodi/index.php/jodi/article/view/197/174

Westell, M. (2006). Institutional repositories: proposed indicators of success. *Library Hi Tech*, *24*(2), 211–226. https://doi.org/10.1108/07378830610669583