PALNI Collaborative Institutional Repository :

Activities and Opportunities

Executive Summary

PALNI's Executive Director, Scholarly Communications Director, and Institutional Repository Task Force have examined closely the IR landscape and platform options for a cost-effective repository suitable for the Private Academic Library Network of Indiana (PALNI) consortium. As a product of this investigation and by pursuing cooperative relationships with other consortia, PALNI has two projects using the platforms Hyku and Islandora. These are the two solutions we've deemed most viable and most closely matching our guiding vision and values. We envision the Islandora project will be ready for production in the FY19 year.

Project Members

Amanda Hurford, PALNI Scholarly Communications Director, is the primary author of this paper, under the direction and support of PALNI Executive Director Kirsten Leonard. PALNI Institutional Repository Task Force members Sudha Anand, Caitlin Balgeman, Brooke Cox, Jennifer Duplaga, Jerry Nugent, and Christa Welty reviewed it and contributed content. Many thanks to all in the PALNI community who reviewed and edited this paper.

FAQ In Brief

- ### What is an Institutional Repository?

  An institutional repository (IR) is a digital collection capturing and preserving the outputs of a single or multi-institution community, in formats such as journal pre-or post-prints, conference presentations, theses, dissertations, and exemplary student works. Digital collection management systems like CONTENTdm are not the same as IRs. See Defining "Institutional Repository.".

- ### Why should PALNI libraries want to host an IR?

  IRs perform a valuable function by creating open access to campus scholarship, and doing so may increase the library's profile on campus and demonstrate its value. Through IRs, PALNI libraries can participate in scholarly communications and the open access movement, as well as help their institution demonstrate its value. See Benefits.

- ## What have we done in the search for a consortial IR?

  We've surveyed PALNI libraries about needs for an consortial IR. Today, only four of PALNI's supported institutions have official IR systems, all using individually operated instances of Digital Commons by bepress. In 2013, 65% of surveyed PALNI schools were interested in participating in a collaborative IR system. We have investigated CONTENTdm, Digital Commons, DSpace, eScholarship, Hyku (formerly Hydra-in-a-Box), Islandora, and an early-phase OCLC prototype as a potential consortial IR solution. See PALNI's IR History.

- ## What are the key aspects of the IR landscape today?

  The acquisition of bepress by Elsevier in July of 2017 has prompted a number of libraries to reconsider whether Digital Commons is a good IR solution. Also in 2017, David Lewis published his popular *2.5% Commitment*, urging libraries to dedicate 2.5% of their budgets to fund open access efforts, including the development of institutional repositories. Open source repository systems are becoming increasingly more attractive than commercial turnkey solutions. Platforms such as academic social networks and disciplinary repositories offer alternatives for faculty to share scholarship, highlighting the urgency for libraries to operate an IR. See IR Landscape Today.

- ## What is PALNI looking for in an IR?

  We want a solution with a sustainable and controllable cost structure that functions as a true IR, promotes open access, and allows for collaborative management. The system should be multi-tenant, open source, future-facing, interoperable, user-centered, user-friendly, flexible, scalable, improvable, customizable, and community-driven. See PALNI IR Vision and Values.

- ## Are there any examples of viable shared IRs?

  We found several examples of viable shared IRs to look to. They use Digital Commons, DSpace, eScholarship, and Islandora.
  See Evaluating the Shared IR Landscape.

- ## What solutions are we pursuing?

  PALNI is leveraging partnerships with peer consortia Pennsylvania Academic Library Consortium, Inc. (PALCI) and Washington Research Library Consortium (WRLC) to pursue opportunities to build shared institutional repositories using Hyku and Islandora. See Partnerships and Solutions.

- ### Why pursue multiple opportunities?

  Hyku and Islandora repository platforms have different foci, communities, and opportunities. Islandora is ready to be deployed now, and WRLC has already optimized their installation. In PALNI's Hyku development, a major difference is that Hyku is focused on multi-tenancy in a way that Islandora is not. Also with Hyku, there is an opportunity to create an IR that will be ultra-low cost to maintain and can scale over several organizations, and be deployed at or across multiple consortia with development. However, Hyku is not ready to deploy now and not likely to be production-ready for a year. By exploring both options, PALNI is able to proactively ensure that we have choices in both the IR and digital collection management software spaces. See Why Both?

- ### What are the financial implications to PALNI and its supported institutions?

  We envision our IR solution to be cost-sustainable. Ultimately, the cost per-institution will be low. This cost will come from PALNI budgeted funds, and we will make every attempt to avoid passing these costs onto our supported institutions. See Financial Implications

- ### When can we expect a PALNI IR to be in place?

  Near the end of FY19, we plan to have the initial Hyku pilot completed as well as our setup of Islandora. We can then begin production in one or both of these platforms, depending on the outcome of these early phases. Once an IR system is in place, PALNI will offer migration support. See Timeline.

- ### What will happen to the PALNI CONTENTdm instance?

  CONTENTdm will be continued, while we will consider whether Islandora or Hyku can manage digital cultural heritage content on parity with CONTENTdm. PALNI will continue to offer CONTENTdm to its supported institutions as long as necessary. See Migration.

Background

Defining "Institutional Repository"

In its 2002 position paper, the Scholarly Publishing and Academic Resources Coalition (SPARC) defined an institutional repository as "digital collections capturing and preserving the intellectual output of a single or multi-university community." Traditionally, institutional repositories focus on scholarly outputs of faculty and students in the form of journal pre-or post-prints, conference presentations, theses, dissertations, and exemplary student works.

Common IR platforms are DSpace, Digital Commons by bepress, and ePrints. They typically focus on content presentation, readability, discoverability, indexing, download stats, and user dashboarding to effectively deliver this content.

In contrast to IRs, digital collection management systems like CONTENTdm are less focused on scholarly outputs and instead often contain institutional records, cultural heritage materials, and items of an artistic nature. The term digital asset management system (DAM) is sometimes used interchangeably to mean digital collection management system, but usually refers to the systems businesses use to organize their multimedia files, often employing digital rights management. Closely related to IRs are subject repositories, research management systems, faculty activity platforms, and digital preservation systems, although they provide distinct and separate functions.

Institutional repositories, while traditionally focusing on text-based scholarship, may house other types of content such as images, which can be effectively separated out using a collection-based structure. Digital collection management systems can house scholarship but many lack the desired functions of an IR (readability, indexing, download stats, and user dashboarding). For the purposes of PALNI's exploration, we'd like to implement a system that effectively does both the functions of an IR and a digital collection management system.

Benefits

Institutional repositories have a number of benefits to both the library and academic campus. Opportunities for direct engagement with faculty and other institutional players are presented by working on an IR, and doing so can increase the library's profile on campus by demonstrating value of the library. IRs collect and openly display institutional scholarship, providing the opportunity for libraries to act as the campus informational hub and serve a role in faculty development. Academic libraries can use an IR to move from an "outside in" model to an "inside out" one that pushes out local content rather than only gathering content from elsewhere. This keeps libraries relevant through changes in the digital information ecosystem. Done well, IRs highlight the librarian skill set, and capitalize on the traditional librarian purview of preserving information and making it accessible to a wide audience.

On a larger scale, through IRs libraries can become increasingly active participants in the overall scholarly communications system, contribute to the open access movement, potentially reduce subscription costs, and demonstrate value and marketability of an institution. SPARC suggests that IRs accomplish the following to address strategic issues facing higher education:

> "Provide a critical component in reforming the system of scholarly communication--a component that expands access to research, reasserts control over scholarship by the academy, increases competition and reduces the monopoly power of journals, and brings economic relief and heightened relevance to the institutions and libraries that

support them; and have the potential to serve as tangible indicators of a university's quality and to demonstrate the scientific, societal, and economic relevance of its research activities, thus increasing the institution's visibility, status, and public value."

PALNI's IR History

In 2013, PALNI's IR Task Force laid significant groundwork toward the ultimate goal of a shared institutional repository system. Their initial recommendation was to tackle an IR investigation in a multistep process that included product evaluation, piloting, and long term support/maintenance. They also administered two surveys, an initial one for directors and another to follow up for a more specific needs assessment.

The group found that over 70% of PALNI schools did not have an IR system in place, and over 65% were interested in participating in a collaborative IR system. Budget constraints were and remain a large factor—the greatest impediment identified in the survey. Following closely behind were staffing limitations and insufficient technological support, highlighting the impact that management of an IR at the consortial level might have.

Of the four of PALNI's supported institutions that have official IR systems, all are using individually operated instances of Digital Commons by bepress. One library recently dropped Digital Commons due to cost and is currently using CONTENTdm as a stop-gap. Several others are using CONTENTdm to display collections of scholarly content, but not in a systematic IR-like way. The desire for a low cost, multi-tenant IR solution for PALNI still exists, and has perhaps intensified in the five years since our initial survey.

The group explored the institutional repository systems DSpace, Digital Commons, Islandora and an OCLC exploration of adapting CONTENTdm to include IR functionality. They also evaluated feasibility of using PALNI's current installation of CONTENTdm for IR workflows. Major roadblocks were found in each solution, such as cost for Digital Commons, technical demands of self-hosting DSpace, and an overall lack of complete specifications for hosted services. Also identified were the numerous ways that CONTENTdm, while largely an effective digital collection management system, is inadequate for the specific workflows of a true IR. CONTENTdm is not set up to efficiently discover and display text, it is not indexed in Google Scholar, and it lacks user functionality desired for an IR. Another overarching hindrance is the lack of multi-tenant functionality in all these systems. OCLC also abandoned the exploration of extending CONTENTdm to include IR functionality.

In 2015, interest for the Hydra-in-a-Box (now known as Hyku) project emerged, which highlights multi-tenancy and straightforward web-based workflows. The IR group then decided that there was not an existing IR tool that would work for PALNI, and recommended that they continue to explore options. While waiting for a suitable system to surface, they proposed that PALNI offer a

suite of publishing services including Open Journal Systems, Open Conference Systems, Open Monograph Press, and the digital collection and exhibit platform Omeka. These platforms are currently installed on PALNI servers and available for use. The Library Publishing Task Force is now examining these platform services and formulating future recommendations for their administration.

In 2016, the IR Task Force recommended Islandora via discoverygarden as the platform for PALNI's shared IR (and digital collection management system) with the knowledge that migration to Hyku would be manageable once that solution was available. This proposal was approved by the board, but not implemented at that time. While Islandora was still a highly desirable system, there were concerns at the PALNI level about adequate capacity within PALNI to manage the project and there were some concerns about the proposal from discoverygarden. More information was needed about discoverygarden and its update strategy. A PALNI solution must retain access and ownership of our data, have the ability to customize and improve, and ensure all data and improvements are portable to another host. Additional research and resources were needed before choosing to move forward, including the creation of the PALNI shared Scholarly Communications Director to oversee the project.

When the Scholarly Communications Director arrived at PALNI in 2017, Hyku had made large strides, and efforts were directed to investigate this platform while continuing to research other options. The beta version of Hyku had just been released, and a pilot of HykuDirect was starting, with wide deployment targeted for the fall. PALNI's local install of Hyku beta was deemed unsuccessful due to our inability to deploy it on the web, but we pursued several options to pilot Hyku with service providers such as Ubiquity Press and Duraspace. Those early efforts did not move forward, but PALNI positioned itself as a strong partner in the shared IR landscape.

IR Landscape Today

The IR landscape is an evolving one, and has undergone significant developments in the past year. The acquisition of bepress by Elsevier in July of 2017 has prompted a number of libraries (including those supported by PALNI) to consider migrating from Digital Commons to another platform, and others to effectively rule it out as a potential IR provider. At roughly the same time David Lewis published his popular *2.5% Commitment*, urging libraries to dedicate 2.5% of their budgets to fund open access efforts, including the development of institutional repositories.

Many institutions and consortia are now working to define options, identify best practices, and in some cases, potential for collaboration in order to add to the open scholarly commons. While requiring a greater technical support base, open source repository systems are becoming increasingly more attractive than commercial turnkey solutions which are expensive and might engage in undesirable practices. Academic social networks like academia.edu are increasing in

popularity with faculty as well as disciplinary repositories like arxiv.org. These platforms offer alternatives for faculty to share scholarship, but highlight the urgency for libraries to offer a central home for campus outputs.

PALNI IR Vision and Values

In consultation with the current IR Task Force and the Executive Director, and after reviewing documentation of past efforts and survey data, the Scholarly Communications Director developed a specifications document for a potential consortial IR. This document, PALNI's Consortial IR Vision provides a graphical sketch of our desired multi-tenant environment, maps out specifications for several functional levels, and defines user management needs. This document has been shared widely across the consortial and library communities and serves a means to easily convey our vision to partners and service providers.

Along the way, we've had the chance to evaluate what is important to PALNI in an IR system, in order to help evaluate opportunities, products, and relationships. The following values have emerged.

PALNI's collaborative institutional repository should be a system which:

- Offers a **sustainable and controllable cost structure** for development, infrastructure, install, maintenance, storage, and support
- Functions as a **true IR**, offering as close as possible to parity with Digital Commons' features
- Is **open source**
- Promotes **open access**
- Allows for a **multi-tenant** environment, in which individual branding and theming at the institutional and consortial level is possible
- Is **collaborative** and can allow cooperative management across teams
- Is **supported** by service providers and/or community developers
- Is **future-facing** and not rooted in old technology structures
- Is **interoperable** and allows **free-flow of data**. Easy import and export of metadata and objects is possible.
- Is **user-centered** and **user-friendly** to users of all types (librarian, faculty, student)
- Is **flexible and scalable** in design and workflow
- Is **improvable and customizable** based on user needs, and developments can be shared back to the community
- Is **community-driven**

Evaluating the Shared IR Landscape

The PALNI IR Task Force was rebooted in 2017 under the direction of the PALNI Scholarly Communications Director. Work of the IR Task Force has focused on evaluating the shared-IR landscape: those IRs that are collaborative in nature either between institutions or campuses of an academic system. A small number of viable platforms for shared IRs emerged from this new investigation: eScholarship, DSpace, and Islandora. There are examples of Digital Commons providing a unified search as means to provide a shared IR (e.g., Iowa Research Commons), but Bepress has not offered a consortial solution.

University of California's eScholarship was a favorite platform in this investigation from a user perspective, offering near parity with Digital Commons. Upon further research, we found that this was a highly customized system out of California Digital Library that would require too much development to be adapted for PALNI use.

Several systems use DSpace as a shared repository provider, such as: California State University, Constellation, Indiana University, ShareOK, Texas Digital Library, University of Missouri Library Systems, University System of Maryland and Affiliated Institutions Library Consortium, and University of Wisconsin. Although widely in use as a shared IR, DSpace is not a viable solution for PALNI. These systems are clean and straightforward in their design, but often lack institution level branding and visual interest, and are known to be inflexible. PALNI consulted with Atmire, one of the major service providers for DSpace about building a multi-tenant version of the application for us, and they preferred to instead build multiple instances, which we've learned is not sustainable from OhioLink's experience. We also learned that many systems using DSpace are actively migrating to another system, or looking to do so.

Islandora is also in use as a shared repository platform by, among others: British Columbia Electronic Library Network (Arca), Colorado Alliance of Research Libraries, Florida Library Virtual Campus, Marmot Library Network, Minnesota PALS, and Washington Research Library Consortium. While operating in a variety of ways, these sites evidence the possibility of a multi-tenant like system that can function as both IR and digital content management system, offering institution level branding and controls with central administration. While largely functioning as a digital collection management system, Islandora has an available solution pack called Islandora Scholar that directly addresses specific IR functionality such as supporting Google Scholar discovery, citation creation, and embargos. Islandora has emerged as the only existing and currently shared repository platform in use that might be viable for a PALNI IR, reinforcing the findings of the earlier iteration of the IR Task Force.

The chart below compares shared-IR options with PALNI's desired IR attributes.

	DSpace	Islandora	eScholarship
Number of Shared IRs Located/Evaluated using this Platform	8	6	1
Support Available	Community /Provider	Community /Provider	Only to UC Community
Open Source	Yes	Yes	Yes
Institution-level Branding	Somewhat	Yes	Somewhat
Modernity of Design	No	Yes	Yes
Easy Data Exchange	Yes	Yes	Unknown
Flexibility of Structure	No	Yes	Unknown
Customizable Interface	Somewhat	Yes	Unknown
Enhanced User Experience	No	Somewhat	Yes
Robust Reporting Analytics	Somewhat	Somewhat	Unknown
Google Scholar Integration	Yes	Yes	Yes

The IR Task Force did not evaluate Hyku in this investigation since it is not currently in use as a shared IR (while in development as such by Ubiquity press for the British Library and now Notch8). Hyku's list of features is available here. This documentation indicates that our desired IR criteria are already met or currently in development for this platform. In addition to pursuing Islandora as a currently viable solution, the PALNI IR Task Force is interested in piloting Hyku when it is developed to the point of viability.

Partnerships and Solutions

Concurrent to the work that the IR group is doing is the research and outreach that the PALNI Executive Director and Scholarly Communications Director have executed in order to engage the consortial community and service providers in conversations about collaborative institutional repositories. We've partnered closely in this research with the Pennsylvania Academic Library Consortium, Inc. (PALCI), a peer consortium in nearly the same position as PALNI in the hunt for a shared IR. In this partnership, we've led discussions at both national (American Libraries Association Midwinter Meeting) and international (International Coalition of Library Consortia North American Meeting) levels, and cultivated an interest group for consortia seeking collaborative institutional repository solutions. Through these discussions, we've made effective connections, gathered needs information, heard about successes to model and failures to avoid, and found creative ways to partner with other actors in this space.

PALNI's efforts to engage peer consortia in the quest for a suitable collaborative IR have resulted in two potential opportunities for IR projects, which focus on strategic partnerships. These projects can capitalize on the value of combined expertise. Also, in exploring multiple opportunities we are afforded the option to compare and evaluate two very different platforms in order to choose the best path for PALNI. Both of these opportunities offer a paradigm shift in how IR services have traditionally been acquired in libraries. Our focus will be on collaboration, community, and sharing, in contrast to a commercially-based vendor model as presented by Elsevier, or the "go it alone approach" of self-hosting and developing open-source solutions in-house.

We've suggested narrowly scoping the pilot phase to focus on developing functionality for hosting electronic theses and dissertations (ETDs) and open educational resources (OERs), while generic and image content types can be uploaded too. This scoping addresses the most pressing need presented by the emerging work of the PALSave: PALNI Affordable Learning initiative. Similarly, as a result of limiting scope, OhioLink has demonstrated success with an ETD-focused project. This would be a good opportunity for PALNI to gain success and ultimately broaden our scope to other materials such as faculty publications.

PALNI Hyku Project with PALCI and Notch8

PALCI has shared information with PALNI about their IR journey. They participated in Duraspace's HykuDirect pilot in 2017 and shared their experience with not only PALNI but the greater community. PALCI and PALNI have fostered a relationship with Duraspace, a key early developer of Hyku. Duraspace has listened closely to the specific use case of consortial IR, and wrote an IMLS grant to pursue this development with PALNI and PALCI as partners. While the grant bid was unsuccessful, the grant application process led us to solidified our needs. Duraspace had begun a Hyku Service Providers meeting/Interest Group. Duraspace distributed our project vision with that group and Notch8 volunteered as a willing developer of Hyku for PALNI and PALCI. Notch8 has successfully worked for the American Theological Library Association (ATLA) on a Samvera project. ATLA is following the developments of Hyku. The PALNI/PALCI Hyku project (CC-IR), will focus on Hyku in the consortial environment, and will benefit the larger community by sharing back the developing and sharing helping to carry forward the potential for a multi-tenant IR specifically for consortial use.

PALNI Islandora Project with Washington Research Library Consortium (WRLC)

WRLC has been an active member of the shared-IR space for some time. Previously they had a shared instance of Dspace, and migrated to Islandora which is more successful. They have offered to host Islandora for PALNI on their servers and offer design and maintenance support for a fee. This partnership will dramatically lower the barrier and costs for PALNI to start using Islandora. Additionally, we have WRLC's technical expertise on our side in order to navigate this

new territory. WRLC's institutions bring significant resources and knowledge and WRLC has optimized Islandora design and workflow. WRLC has pursued this partnership to bring PALNI's talents and power to Islandora development. This project allows PALNI to implement a proven open-source IR solution in use at WRLC, while also developing the cutting edge Hyku software with PALCI and Notch8.

Why Both?

Hyku and Islandora repository platforms have different foci, communities, and opportunities. Islandora is ready to be deployed now, and WRLC has a successful consortial-focused installation. Partnering with them provides multiple advantages because WRLC has more capacity to hire and direct development support. They already have good template designs in place. Through this partnership PALNI also has the opportunity to further develop Islandora with an organization that has strong interest and capacity to do so. WRLC wants PALNI to contribute to the development of Islandora as a community-driven platform. PALNI will be able to benefit from this optimized environment and aggregation of resources for many years.

PALNI's Hyku partnership with PALCI and developer Notch8 is focused on the major difference between Islandora and Hyku, multi-tenancy. Hyku is focused on multi-tenancy in a way that Islandora is not. This native capacity to host multiple, separate user-facing applications is critical for a cost-effective shared IR environment. Workflows across institutions can be more easily developed in Hyku over Islandora. However, Hyku is not ready to deploy now and not likely to be production-ready for a year. With Hyku, there is a development opportunity to create an IR that will be ultra-low cost to maintain and can scale over several organizations, and be deployed at and across multiple consortia. Our partnership's initial focus on developing out workflows for ETDs and OER material can create a vital hub for this original campus-based content, and in doing so raise the profile for our supported institutions.

By exploring both options, PALNI is able to ensure that we have choices in both the IR and digital collection management software spaces while doubly contributing to the open source and open access communities in a practical, innovative, and collaborative fashion. Additionally, these two partnerships and tools allow us to develop tools that meet divergent needs and to also create high profile data centers.

Financial Implications

PALNI recognizes financial pressures facing its supported institutions and offers these pilots in response to stated needs and priorities while keeping an eye on long term budget implications. We envision our IR solution to be cost-sustainable. By choosing non-commercial solutions we are working to insulate ourselves from the risk of cost-hikes and forced migrations. Wherever possible, we'll utilize in-house resources and expertise in order to keep costs low. The goal is to develop a partnership and platform that produces a low cost per-institution, significantly reduced

from what would be required if institutions were to install a system individually. PALNI is funding the program through the PALNI central budget, and will make every attempt to avoid passing these costs on to the institutions.

Timeline

PALNI is pursuing opportunities to install Islandora and develop Hyku simultaneously in FY19. A work agreement is signed with Notch8 which includes a phased approach. Phase 1 is setting up a Hyku application, some work on workflow and permissions, and basic theming and branding, which will be covered by the $18,000 budget set aside for IR development in FY18. The budget funds for Phase 2 will be provided by PALCI. Our intention is to complete this phase in the first half of FY19 and begin testing.

We are now working with WRLC to create a project plan and set a timeline to for deployment of the PALNI Islandora shared IR with an intention of completing testing over summer 2019.

Near the end of FY19, we plan to have the initial Hyku pilot completed as well as our setup of Islandora. We can then begin production in one or both of these platforms, depending on the outcome of these early phases. The Scholarly Communications Director will continue to monitor the shared IR landscape and review new solutions on an ongoing basis.

Migration

Once an IR system is in place, PALNI will offer migration support to institutions wishing to move collections from Digital Commons and/or CONTENTdm via PDAT and the eventual team responsible for administering the platform(s). Some collections containing institutional scholarship may be redundantly migrated to Hyku and/or Islandora during FY19 as we explore functionality of those systems. While we will consider using Islandora and Hyku to manage digital cultural heritage content, neither will act as an immediate replacement for CONTENTdm. PALNI will continue to offer CONTENTdm to its supported institutions as long as needed.

Resources and More Info:

The Case for Institutional Repositories: A SPARC Position Paper

Discoverygarden Islandora overview

Hyku Features

eScholarship: University of California

Institutional repository software comparison

Institutional repository software comparison: DSpace, EPrints, Digital Commons, Islandora and Hydra

Institutional repositories: exploration of costs and value

Islandora