Building Hyku for Consortia: Cross-Consortial Partnerships for Community Infrastructure

Kirsten Leonard, PALNI Executive Director & Jill Morris, PALCI Executive Director
CNI Project Update
December 3, 2020
Outline

1. About Us
2. Need and Vision for Collaboration
3. Project Specifications and Development Efforts
4. Future Plans
5. Call to Action
Staffing the Hyku for Consortia Project

Kirsten Leonard and Jill Morris, Project Directors and Principal Investigators

Amanda Hurford, PALNI, Hyku for Consortia Product Owner, co Project-Lead

Gretchen Gueguen, PALCI, Hyku for Consortia Project Manager, co Project-Lead

Nicholas Stanton-Roark, PALNI, Pilot liaison, Documentation and QA Testing

In addition to us:

Notch8: Software Development

Product Management Team: Design and Requirements
Thank you to IMLS

PALCI and PALNI were awarded $172k from IMLS in June 2019 to support a second development phase (the first was funded by PALNI) to establish a multi-tenant, consortial IR application, and to explore the potential for supporting shared IR services infrastructure across two consortia.
Collaboration at our Core
The “PALs”

**Who?**
- 71 public and private academic and research libraries, incl. 6 ARLs: NYU, Penn State, Rutgers, Temple, UPenn, UPitt

**Where?**
- PA, NJ, NY, WV

**Structure?**
- 501(c)(3) membership organization

**Staffing?**
- 4 FTE Staff

---

**Groups serving diverse membership:**

*Problem:* finding common ground

---

**Groups serving specific institution types:**

*Problem:* prioritizing where to act
Collaborating to Address Barriers & Shared Challenges

- Need to control costs
- Market consolidations and acquisitions
- Limited staffing capacity, expertise, and tech resources
- Need for easy-to-use open source alternatives
- Importance of discovering Open Access material
- Need for innovation AND scale
- Desire to share the work
- Desire for flexible systems to meet diverse needs

Hyku for Consortia Vision & Values - Partnering in order to:

1) Contain and share costs
2) Build IR efficiencies leveraging shared expertise, resources, tech management, and scale
3) Develop an open source community infrastructure for innovation and interoperability
4) Hit the sweet spot in the marketplace, with a “good enough” technology solution to
What is Hyku?
Hyku is a standardized installation of the open source digital repository Hyrax meant to meet the largest possible set of basic repository needs.

Hyku has a predetermined set of functions, types of objects and metadata out of the box. It can also run as a multi-tenant solution.

Hyku can be installed and configured as a single application, instead of needing to configure the various components as you would with Hyrax.
Multi-Tenant
PALs working together

Methodology:
- Gathering input from many libraries
- Synthesize into prioritized needs
- Translate to the Hyrax/Hyku environment

Phase 1: 2017-2018
Get Hyku up and running

Phase 2: 2018-2021
Improve features for multi-tenant administration

Phase 3: 2021-2023?
Remove barriers to adoption for institutions
Hyku for Consortia Development

Grant Goals and Deliverables

- Collaborative workflow support and bulk uploads  Development in progress
- Work types w/metadata support for OER and ETD  Development in progress
- Themed templates  Specifications gathered
- DOI integration  Gathering requirements
- Cross-tenant searching  Gathering
- Multi-tenant shared works  Gathering requirements
- Contributing code / integrating community development  In progress
# Adding Collaborative Workflows

<table>
<thead>
<tr>
<th>Scenario name</th>
<th>Scenario description</th>
<th>What is the resource type?</th>
<th>Where does the content come from?</th>
<th>Who uploads?</th>
<th>Who describe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty self-submit</td>
<td>a faculty member uploads their own work</td>
<td>pdf</td>
<td>directly from faculty member</td>
<td>faculty member</td>
<td>faculty member, library staff, data, library staff, data reuse</td>
</tr>
<tr>
<td>Archival collection</td>
<td>Archives has a collection they want to put online in connection with the library and a</td>
<td>images/ text documents (possibly need full-text searching)</td>
<td>Faculty or archives select content, library digitizes</td>
<td>Library staff to an area that is only viewable by staff involved</td>
<td>Small data, library staff</td>
</tr>
<tr>
<td>Library managed</td>
<td>Library staff collect and upload materials received from faculty</td>
<td>Document, link, or metadata only record</td>
<td>Faculty or staff via a web submission form, direct email to library staff, library staff productively searching and adding materials</td>
<td>Library staff</td>
<td>Library staff</td>
</tr>
<tr>
<td>Journal hosted on IR</td>
<td>Student, faculty, and staff journal submissions</td>
<td>Document</td>
<td>Journal authors (students, faculty, staff, outside submitters)</td>
<td>Journal editors</td>
<td>Journal editors, library staff</td>
</tr>
<tr>
<td>Theeses and Dissertations</td>
<td>Accepted degree research from institutional department</td>
<td>PDF, but also data objects. images, text, datasets</td>
<td>Authors via a web submission form</td>
<td>Department once approved</td>
<td>Author provides metadata, any control data</td>
</tr>
<tr>
<td>Theeses</td>
<td>Thesis accepted by department chair or thesis advisor</td>
<td>born digital or scanned pdf; may include scores, videos, and/or audio files(native streaming would be wonderful)</td>
<td>Authors via submission form or library collection (for older theses)</td>
<td>Library staff</td>
<td>Library staff</td>
</tr>
<tr>
<td>Journal hosted on IR</td>
<td>Student, faculty, and staff journal submission pdfs, word docs, text files</td>
<td>authors* (see note)</td>
<td>Authors upload original version; library staff uploads final version</td>
<td>Authors provide additional information, library staff verify data</td>
<td>Author provides additional information, library staff verifies data</td>
</tr>
</tbody>
</table>
Project Timeline Overview

● Project initiation and planning: July 2019
● Requirements and design process (iterative): November 2019 - February 2020 (on hiatus)
  ○ Cross-consortial Product Management Team
  ○ Specifications for worktypes, collaborative workflows, and theme templates gathered
● Development sprints (iterative): April 2020 - present
  ○ OER worktype done, ETD worktype is next
  ○ Refining user permissions, roles, and groups for multi-tenant administration
  ○ Prerequisite bug fixes, Bulkrax, ability to hide tenants, feature flexibility
● Testing, Pilot and Feedback (iterative): July 2020 - present
  ○ Production pilots started with 14 institutions (7 each)
● Governance and business model development: ongoing
● Documentation and closeout: ongoing - December 2021
  ○ Identify funding for future development phases
Business Model for Sharing Infrastructure

Vision: PALCI and PALNI will jointly offer IR services desired by member libraries at greatly reduced costs to support a variety of consortial and individual library use cases, with a particular focus on collaboration for general IR services, affordable learning materials, and ETDs.

3 Scenarios:  ★★★  1) Deep Collaboration:  
Shared Administration, Management and Expertise

★★  2) One Consortium as Service Provider: Some cost sharing

★   3) Hosted Solution only; Commercial Service Provider
IR Cost Factors

AWS hosting / server

Support & maintenance

Software development

Staffing / Management

# of institutions

---

**Scenario 1 - Sample Cost Share Models (Rough estimates only, estimated conservatively)**

<table>
<thead>
<tr>
<th></th>
<th>PALCI</th>
<th>PALNI</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AWS / Server Costs</strong></td>
<td>$12,000 (% use * $26,000)</td>
<td>$12,000 (% use * $24,000)</td>
<td>Split based on actual use of storage (used 50/50 split for estimating purposes)</td>
</tr>
<tr>
<td><strong>Support &amp; Maintenance</strong></td>
<td>$6,000</td>
<td>$6,000</td>
<td>Split evenly up to retainer w/NotchB or other contracted support</td>
</tr>
<tr>
<td><strong>Development</strong></td>
<td>$15,000</td>
<td>$15,000</td>
<td>Split evenly up to retainer</td>
</tr>
<tr>
<td><strong>Staffing @ PALCI</strong></td>
<td>$40,000</td>
<td>$40,000</td>
<td>Demand set on salaries + benefits, 2 Part-time or 1 Full-time</td>
</tr>
<tr>
<td><strong>Total Costs</strong></td>
<td>$79,000</td>
<td>$73,000</td>
<td></td>
</tr>
<tr>
<td><strong>Funding Models</strong></td>
<td>Cost recovery</td>
<td>Cost recovery</td>
<td></td>
</tr>
<tr>
<td><strong># Institutions</strong></td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Estimated cost per institution</strong></td>
<td>$3,500</td>
<td>$3,500</td>
<td>Needed to figure out the true cost per institution with cost info based on terms/storage capacity in Hyku so that we can determine the best scaling efficiency. It's possible that if PALCI were to offer Hyku to 30 members that we could take on more of the costs, but does that raise the other costs of support/hosting significantly?</td>
</tr>
</tbody>
</table>
Opportunities for Future Development: Addressing Barriers to Adoption

- Configurable metadata fields per tenant
- Customization of controlled vocab
- Google Scholar indexing

- Sophisticated reporting and metrics
- Faculty profiles
- Integrated journal creation workflows
Rethinking Institutional Repositories

The 2017 CNI *Rethinking Institutional Repositories* report identified a number of major challenges in the IR space and a great diversity of institutional needs.

This is important because a great deal of IR development has happened / is happening in silos;

If we worked together and leverage the diversity of our institutions better, open source systems like Hyku, have the potential to provide previously unavailable scale and greater access to underserved / less well-resourced communities.
Call to Action: Join the Hyku Community

Investments and library commitments to development of open and community-owned infrastructure is needed to contain costs and create viable open source alternatives.

Innovation is needed; Development experience is needed; Diversity is needed.

We believe our partnership is a model for sharing infrastructure, including software + expertise and human resources.

Ask us about a demo for your group of libraries.
THANK YOU!

Questions?

Hyku Demo:  
https://vimeo.com/424366097/1669194f8f

Visit: hykuforconsortia.org

Twitter: @hyku4consortia

Email:

gretchen@palci.org; jill@palci.org;
amanda@palni.edu;
kleonard@palni.edu