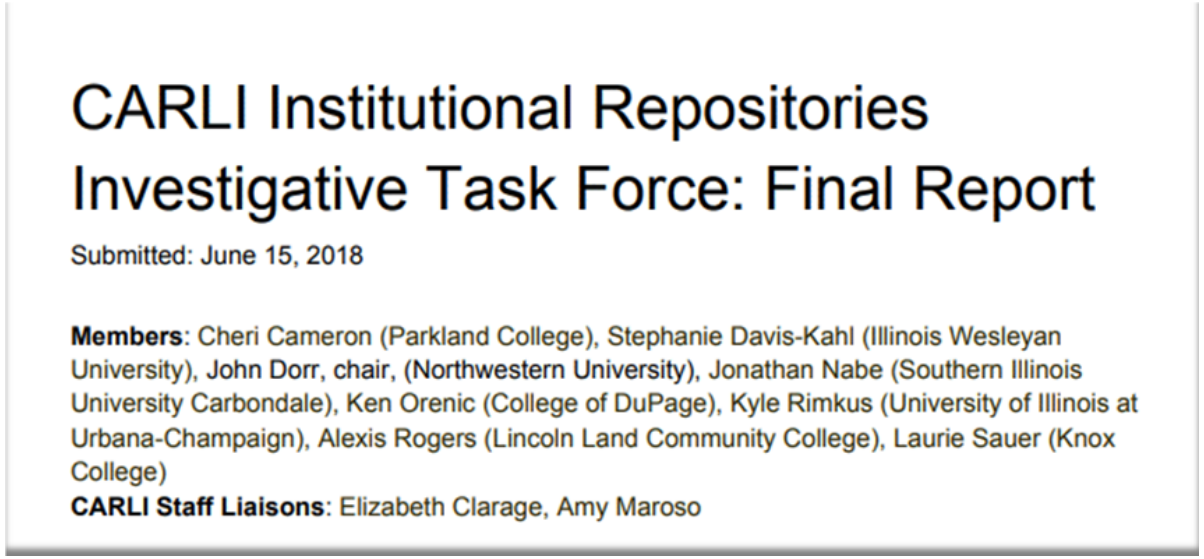
BUILDING A SUSTAINABLE STATEWIDE INSTITUTIONAL REPOSITORY SERVICE

Kyle Rimkus
Preservation Librarian, University of Illinois at Urbana-Champaign
CARLI members Meeting | November 13, 2020

Some slides adapted from presentations by Ayla Kenfield, Christopher Prom, Nicole Rhodes, Seth Robbins, and Joe Troy.

WHAT IS AN INSTITUTIONAL REPOSITORY?

https://www.carli.illinois.edu/sites/files/governance/IRTF_Final_Report.pdf



CARLI Institutional Repositories Investigative Task Force: Final Report

Submitted: June 15, 2018

Members: Cheri Cameron (Parkland College), Stephanie Davis-Kahl (Illinois Wesleyan University), John Dorr, chair, (Northwestern University), Jonathan Nabe (Southern Illinois University Carbondale), Ken Orenic (College of DuPage), Kyle Rimkus (University of Illinois at Urbana-Champaign), Alexis Rogers (Lincoln Land Community College), Laurie Sauer (Knox College)

CARLI Staff Liaisons: Elizabeth Clarage, Amy Maroso

"CARLI defines an IR as a set of services and technologies for the local management and dissemination of digital materials created by member institutions and their communities. Traditionally, IR platforms allow for the self-deposit of content from an institution's community of users. As such, a consortial IR would fulfill a need distinctly different from that of CARLI's CONTENTdm-driven CARLI Digital Collections service."

WHAT DO PEOPLE STORE IN AN IR?

Speaking for UIUC, our IDEALS IR primarily houses research and scholarship produced at our university. This includes scholarly, research-oriented, educational, creative, or other intellectual output of the university community.

- Research papers
- Conference presentations
- Theses and dissertations
- Websites
- Videos
- Soil surveys
- Books
- Interviews
- Code
- Spreadsheets
- Fiction
- Annual reports
- Podcasts
- Maps

INSTITUTIONAL REPOSITORY TASK FORCE (2017-2018)

Charge:

1. Conduct an environmental scan, particularly of other multi-institution-based IRs;
2. Assess the existing IRs (platforms, inventories, and institutional guidelines) within the CARLI membership members' repository structure and inventories;
3. Investigate available software platforms, both open source and proprietary;
4. Investigate migration issues that could arise moving current standalone IRs into a consortial setting;
5. Determine if a consortial IR is feasible and what the structure of the IR might be; and
6. Determine costs associated with the building and maintaining of an IR including what is necessary centrally and from members to construct an effective service.

THE LANDSCAPE OF IR SOLUTIONS

The task force described IR solutions in a period of transition, citing:

- **Uncertainty around bepress Digital Commons.**
- **Uncertainty around the future of DSpace.**
- **The immaturity of Fedora-driven solutions.**
- **The emergence of additional options.**
- **The push into cloud-driven architectures**

They also found that:

- There is significant interest in pursuing a consortial institutional repository among CARLI members.
- There is no evidence of activity in the vendor community around consortial pricing models and the options for building and maintaining a consortial IR within the CARLI organization, as limited by current commitments and staffing.
- The communities building platforms are also intrigued with the idea of developing an infrastructure that could support a consortial IR, but that development is not on the short-term deliverables of any active roadmap.

LOOKING AHEAD TO 2021 BACK IN 2018

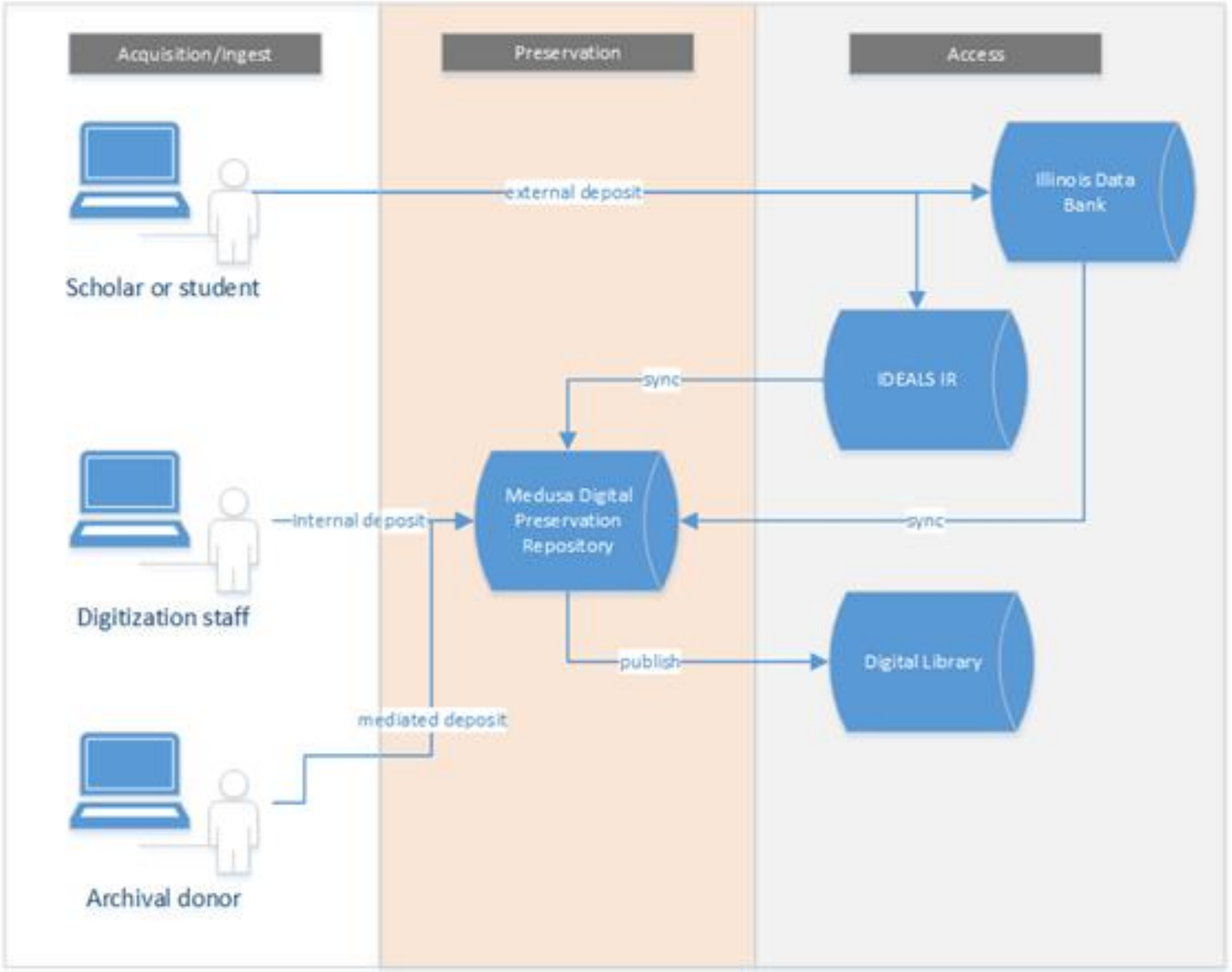
"Taken together, these factors suggest that the optimal strategy a consortium like CARLI ought to take in deploying a consortial IR may become much clearer in two to three years time than it is right now."

"We are confident that in 2-3 years the market will be in a better place to support this type of activity, possibly within the state itself, as the University of Illinois at Urbana-Champaign is actively exploring rebuilding its own IR services with an eye to a multi-tenancy."

INTEGRATING PRESERVATION AND ACCESS

At UIUC, we build access systems on top of our Medusa digital preservation repository, meaning that access and preservation are managed together rather than separately.

Everything deposited into public-facing repository services like IDEALS (our IR) and the Illinois Data Bank (for research data) are backed up in Medusa.



Current paths for preservation and access (simplified view).

A LONG-TERM APPROACH TO BUILDING SYSTEM ARCHITECTURE

Phase I (2012-2015)

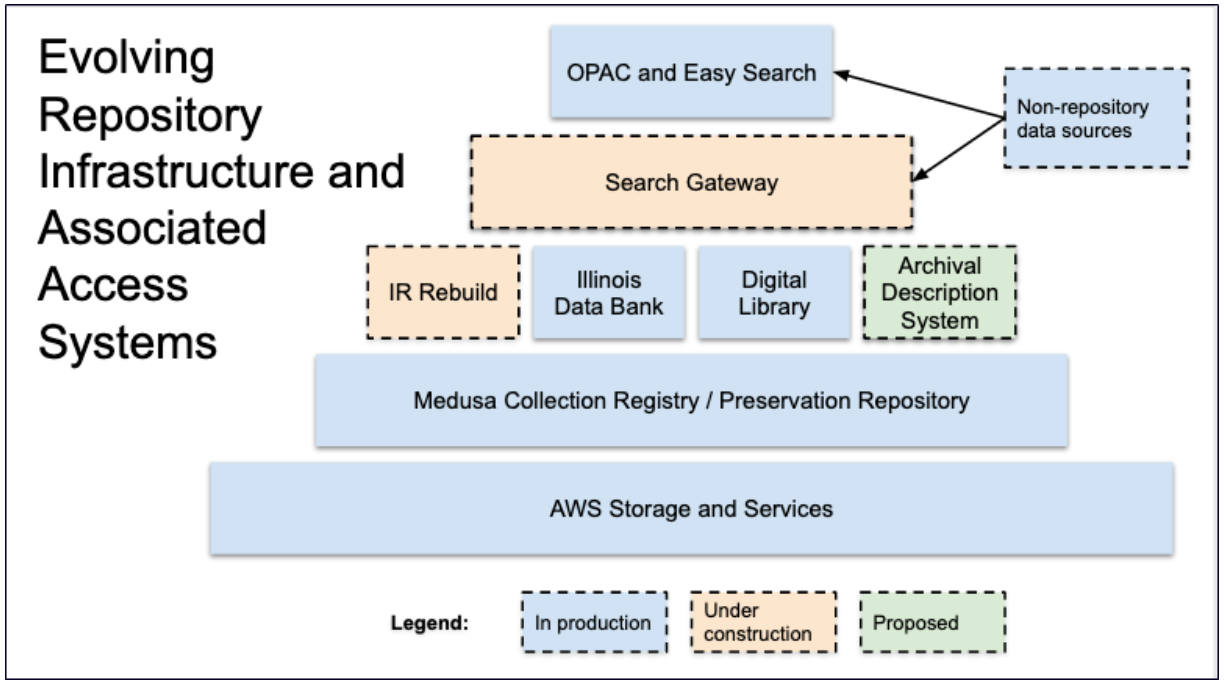
- Build Medusa Preservation Repository (https://medusa.library.illinois.edu/)
- Describe and ingest content into Medusa

Phase II (2015-2018)

- Build access systems on top of Medusa
- Digital collections (ContentDM replacement, https://digital.library.illinois.edu/))
- Illinois Data Bank (https://databank.illinois.edu/)

Phase III (2018-2021)

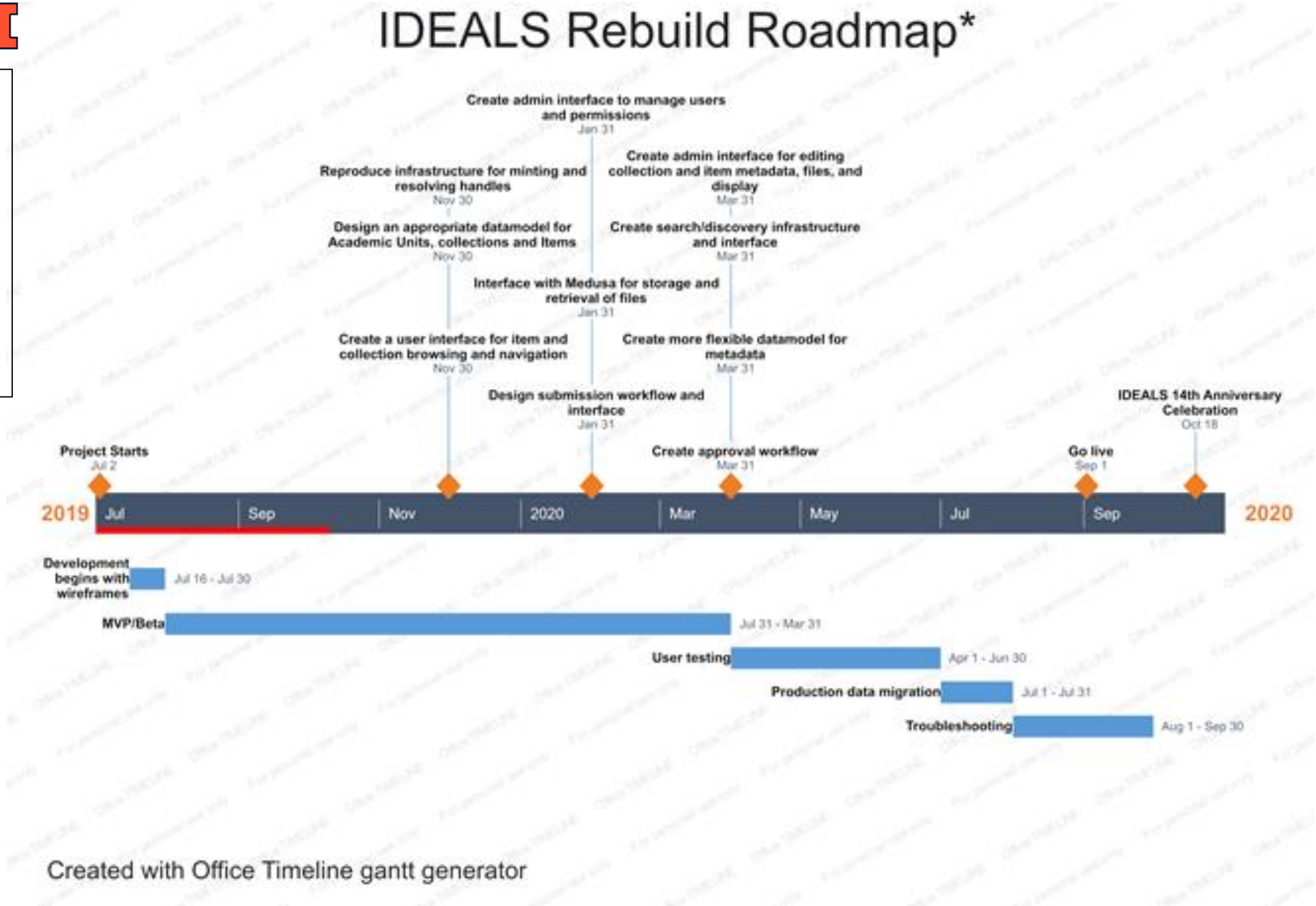
- Shift repository infrastructure into Amazon Web Services
- Build metadata search and discovery service for all local digital collections
- Build IR system to replace local instance of Dspace (underway)



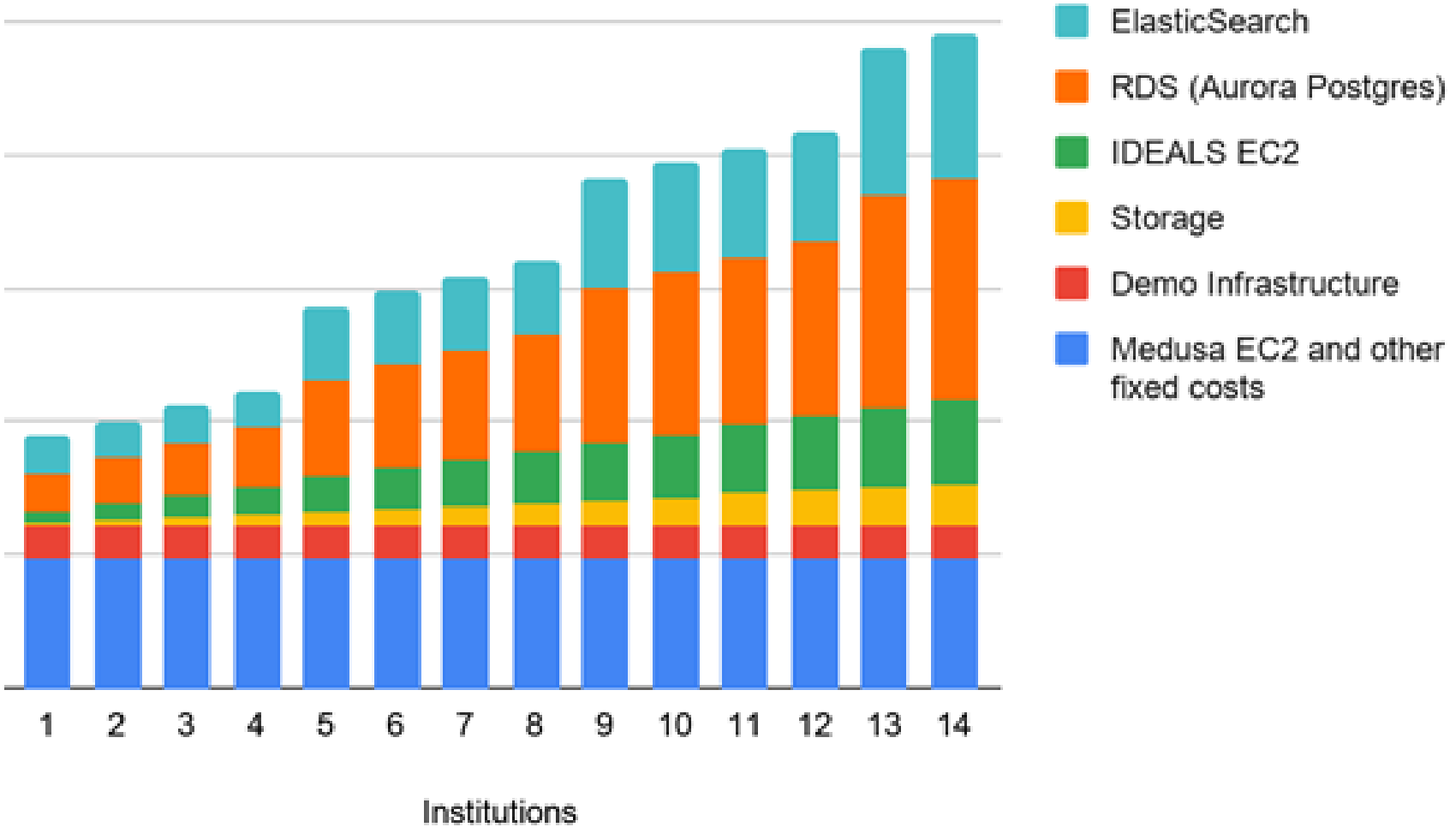
Next steps

- Expand services to allow for multitenancy (access for other institutions)
- Replace Archon archival description system for fully integrated content management suite

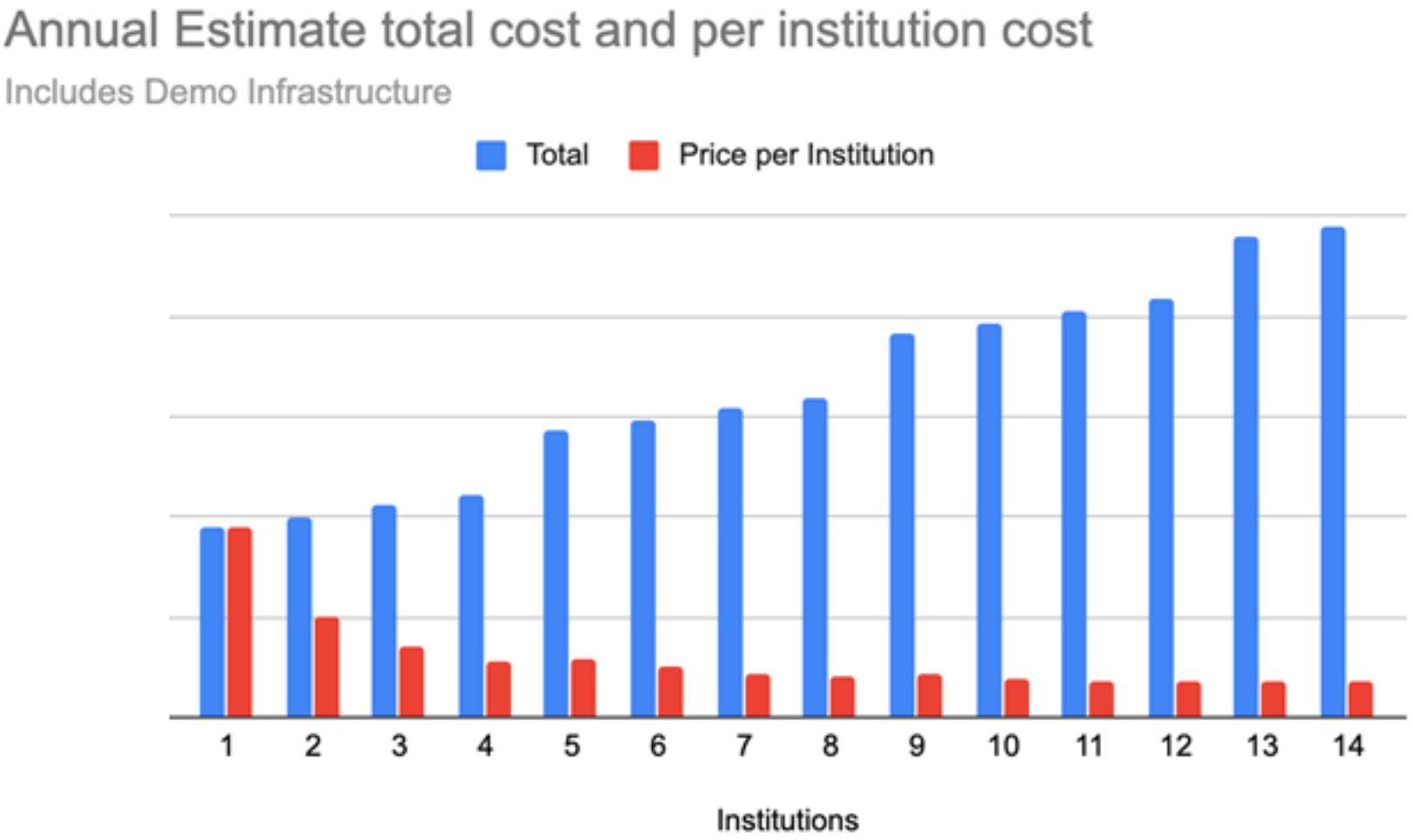
- UIUC IDEALS IR can be found at https://www.ideals.illinois.edu/
- Active since 2006
- Features over 104,000 items



RELATIVE INFRASTRUCTURE COSTS



HOW THIS BECOMES AFFORDABLE AT SCALE



MEMORANDUM OF UNDERSTANDING (UNDERWAY)

Library Responsibilities

a. Collaborate with CARLI to provide system development.

b. Provide ongoing system development, deployment, and maintenance.

c. Provide persistent storage of digital items for preservation and access.

d. Additional technical support may be provided upon request and when mutually agreed upon by both parties.

CARLI Responsibilities

a. Provide user training in the form of introduction to use of State Institutional Repository, how to access and maintain State Institutional Repository.

b. Provide customer support as the primary user contact for questions and troubleshooting.

c. In consultation with Library, develop use policies to share with users.

d. In consultation with Library, develop procedures to share with users.

e. Collaborate with Library to provide system development.

IRTF SURVEY'S TEN MOST DESIRED FEATURES

Feature	Will our system provide it?
Usage Statistics	Yes
Customizable metadata	Yes
Unlimited storage	Yes.
Institutional branding	Yes, in a limited fashion (such as a customizable banner)
Display media formats (audio, visual)	Not really. Media files will be available for deposit/download, but we do not expect to have a powerful native media player
Integration into campus authorization/authentication systems	We don't know; this will require further exploration into member expectations and technical requirements
Suppress content from public view (i.e. campus-only viewing)	Yes
Search engine optimization (SEO)	Possibly; we are looking into it
Batch upload	Yes
Electronic Theses and Dissertations (ETD) workflow	Not a full-fledged workflow manager, although the system can certainly host ETDs

WHAT ARE OUR NEXT STEPS?

- UIUC will complete IR migration, hoping to go live September 2021
- UIUC and CARLI will continue work on planning for cost models, membership policies, etc.
- UIUC and CARLI will solicit participation from potential users of this service in a pilot for late 2021
- If all goes well, a consortial IR would be available at some point in 2022

MIGRATION CHALLENGES FROM BEPRESS AND DSPACE

- Respecting time-based embargoes placed on content in either system.
- Respecting local access restrictions placed on content in the either system.
- For DSpace, respecting local permissions related to which users can and cannot access or edit content filed under specific DSpace “communities.”
- For DSpace, understanding local changes or modifications made in local DSpace implementations that could affect migration plans.
- For bepress, understanding to what extent add-on features such as journal or conference management affect migration of IR content of each respective institution.
- Mapping persistent URLs from the source system to the new one.

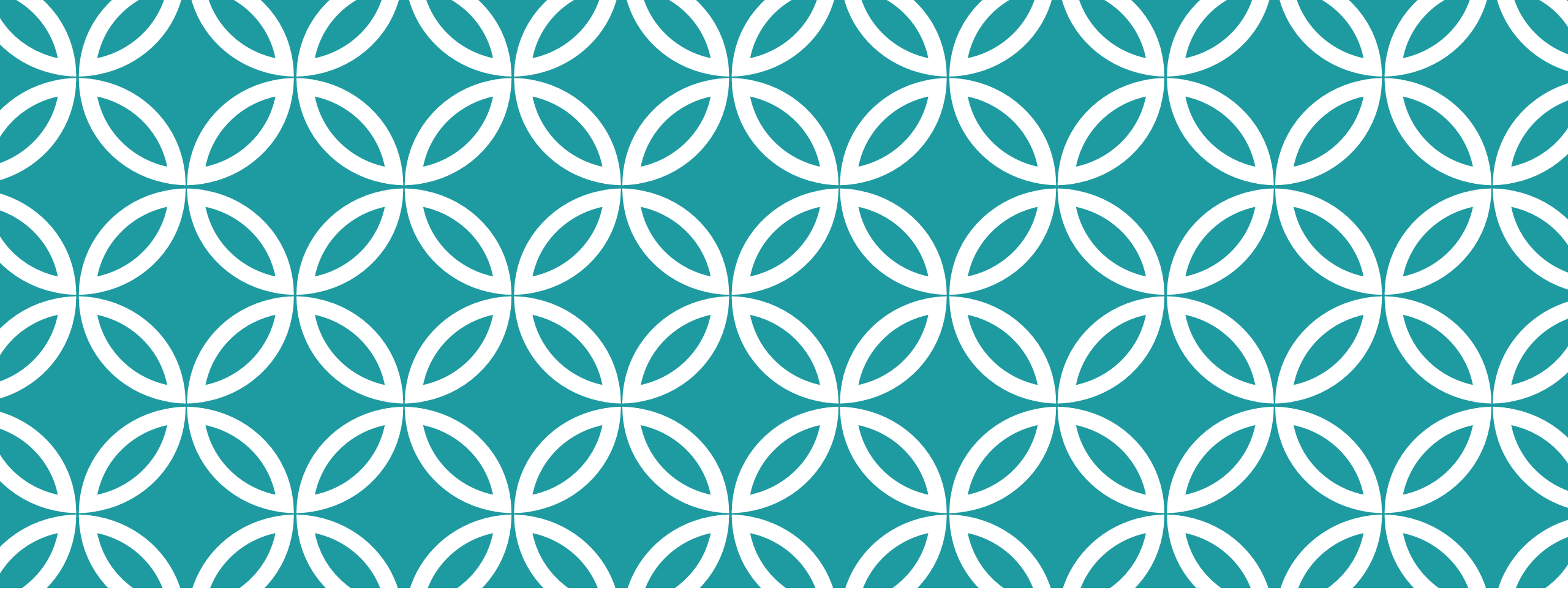
IR PLATFORMS IN ILLINOIS CIRCA 2018

Environmental Scan of IR platforms in Illinois (36 identified)

Platform	Number of IL IRs
ContentDM	1
Digital Commons	21
DSpace	11
Hyrax	1
Homebrew or other	2

Survey Response: Satisfaction with current repository platform (27 responses from members with IRs)

Platform	No. of installations	Very satisfied	Somewhat satisfied	Not satisfied
bepress/Digital Commons	14	8	6	0
DSpace	10	1	7	2
CONTENTdm	1	0	0	1
Other (Hyrax [Samvera]; unknown)	2	0	1	1



THANK YOU!

Kyle Rimkus, rimkus@Illinois.edu