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

Executive Summary 2

History of Group 3

Survey and Results 4

Migration concerns 6

Analysis 8

Recommendation 9

Appendix I: Full Survey Results 11

Appendix II: CARLI Institutions with IRs 27

Appendix III: Meeting Summaries 29

Appendix IV: IR Development at Illinois 30

Executive Summary

CARLI convened an Institutional Repositories Investigative Task Force (IRTF) in October 2017 to investigate the possibility of establishing a consortial institutional repository (IR) in Illinois. After consulting Clifford Lynch's seminal paper on IRs[1] and the Repositories Support Project for model language[2], the Task Force agreed on the following to clarify its approach: "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."

The IRTF then conducted a survey of current IR use among CARLI members. Overall, responses suggested a strong interest in a consortial IR service, underscoring a lack of funding, staff, and technical expertise at numerous institutions throughout the state that currently do not have an IR. In fact, 58% of respondents answered in the affirmative to a question about willingness to participate in a consortial solution (30% of respondents to the same question expressed some hesitancy, due to possible costs and already established IRs). While the survey and its results are presented in detail in the report below, its responses have provided CARLI with evidence of interest in a consortial IR throughout the state, a comprehensive overview of membership concerns, as well as information about the types of content CARLI members would deposit in a repository, and the set of desired features and functions in a potential future IR.

Independent of its survey, the IRTF explored the current state of the IR landscape. The acquisition of bepress by Elsevier has catalyzed many discussions about the possibility of changes to bepress products, contracts, and costs, and uncertainty about its future. While the acquisition is recent, there are efforts underway at some institutions to investigate alternatives to bepress, either individually or as a group. In one such effort, the Scholarly Publishing and Academic Resources Coalition (SPARC) included a key priority in their 2018 Program Plan to advocate for community-controlled infrastructure including "accelerating work on creating new models of repositories."[3] Open source repository solutions are also growing. While DSpace has been a major platform for more than a decade, other options are now available, such as Fedora-driven solutions like Islandora and Hyku, and other locally developed solutions. In addition, the University of Illinois at Urbana-Champaign is developing its own IR solution to replace the DSpace platform on which its own IR operates, and is building its next-generation IR service with an eye to use by other institutions. Overall, the state of IR solutions is in flux, with expectation of significant changes in the coming two to three years.

Based on the general interest in Illinois in a consortial IR solution and the current uncertainty around IR solutions, the IRTF recommends that CARLI continue to support the ongoing effort of this Task Force (either in the same form or with a modified membership and charge) to continue

[1] See Lynch, Clifford A. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." *Association of Research Libraries ARL Full*, https://www.cni.org/wp-content/uploads/2003/02/arl-br-226-Lynch-IRs-2003.pdf

[2] http://www.rsp.ac.uk/

[3] https://sparcopen.org/who-we-are/program-plan/

to monitor the landscape of IR technology, with the expectation that this will mature in the coming years in a way that will better support CARLI members' interests in a consortial IR solution.

History of Group

The IRTF consisted of members Paul Blobaum (Governors State University), 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), and CARLI staff liaisons Elizabeth Clarage and Amy Maroso. The membership changed slightly over the course of the year. It was charged with exploring whether CARLI should operate an IR on behalf of its member institutions. Specifically, as part of its investigation, the Task Force was asked to:

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

In October 2017, CARLI identified and invited library staff from nine libraries to serve on the Institutional Repositories Investigative Task Force. Membership (see above) included representatives from community colleges, 4-year, and graduate institutions. The Task Force convened seven times, either in person or by phone. The group issued this final report in June 2018. A summary of activities that occurred during its meetings is available in Appendix III. Specifically, the Task Force was instructed that one or more of several recommendations was to result from its efforts, based on the outcome of its investigation, and provided the following list of options (these are addressed in the Recommendation section of this report):

1. CARLI does not act as an IR provider for CARLI members at this time.
2. CARLI acts as an IR provider for CARLI members.
3. CARLI acts as a clearinghouse for information and education about IRs but not provide IR functionality for CARLI members.
4. CARLI does not act as an IR provider, but helps to facilitate (e.g., provide meeting space, an email list, announcements) member libraries to pursue shared IR opportunities.
5. CARLI acts as the procurement agent for an IR and offers the product to the members similar to brokered electronic resources.
6. Specifications in a Memorandum of Understanding to which a group could respond, in which CARLI would act as a broker. As an example of such an arrangement in a library consortium, see https://geo.btaa.org/about.

Survey and Results

The Task Force surveyed CARLI member institutions early in 2018 to: determine current use of and satisfaction with IRs; uncover reasons why IRs have not been implemented; and interest in and anticipated use of a CARLI consortial IR. What follows is a summary of survey responses. A full accounting of them is available in Appendix I.

Survey response rate was 55%. Doctoral and Masters level institutions were somewhat overrepresented in the respondent pool; Baccalaureate, Community College, and Special Library institutions were slightly underrepresented (see Appendix I, Table 1).

Only 38% of survey respondents (n=73) indicated that their institutions have an IR. For those responding (n=27), bepress is the most prevalent platform (14 installations) and Drupal is the next most used platform (10 installations). Satisfaction with bepress is higher overall than with Drupal (see Appendix I, Table 2).

Institutions that do not have an IR were asked why they do not have one. The reason chosen by most respondents was “lack of funding to support ongoing storage and maintenance”. The reasons “lack of staff support” and “lack of technical expertise”, similar in nature, both garnered high numbers of responses. Another dominant reason given was that an IR is not a priority. Note: Respondents were able to choose more than one reason. (See Appendix 1, Table 4.) Some comments revealed in more detail overlap between “lack of funding” and “competing priorities”:

- “No stated interest … Also, it would require a commitment in terms of funding and staff that's probably not available in our current budget environment.”
- “Given staffing and budget reductions, we haven't been able to make an IR a priority.”
- “At this point, we are so short-staffed that creating an IR has fallen to the bottom of our list.”

To gauge interest among CARLI libraries for a consortial IR, the survey asked, “If CARLI were to offer a consortial IR solution, would you consider a cost-sharing arrangement with CARLI for your organization to benefit from access to a shared platform and training opportunities?”. Support among respondents was high: 58% answered “yes” while only 12% answered “no”. Some uncertainty exists, however, with 30% of respondents expressing uncertainty. Certainty about adoption of a consortial IR is higher among Baccalaureate and Masters institutions than the other institution types. (See Appendix 1, Table 5.)

For the respondents who indicated they were **not** interested in a consortial IR, respondents were asked to state **why** they were not interested (see Appendix 1, Table 10). The responses were in free text, but some clearly identifiable reasons were given. Cost was the top factor, cited by 10 respondents; this mirrors the above discussion on why institutions do not already have an IR. Several respondents (8) cited the fact that they already have an IR, possibly indicating that a future consortial IR may not supplant the local installation. A smaller number of respondents (5) indicated they would have no need for a consortial IR.

Respondents were asked to **prioritize functions of a consortial IR** (n=38), ranking eight functions against each other (see Appendix I, Table 6). Providing access to student work is the

top priority, followed by providing access to faculty publications and digitized collections. Medium priority functions are having the IR act as a digital preservation service for institutional content and provide access to faculty research data. Functions seen as the lowest priority include: provide access to campus media content; showcase bibliographies of faculty work; comply with government mandates for open access to research.

Respondents were asked to determine the type of content that they may put in a consortial IR. (See Appendix I, Tables 3 and 7.) These responses correlate generally with the type of content that is already housed in IRs of individual CARLI libraries. The top five content types garnering 50% or more of responses are:

Type of content potentially in consortial IR	**Type of content in existing CARLI libraries' IRs**
Rank 1: Archival content	Rank 1: Student-generated content
Rank 2: Scholarly publications	Rank 2: Scholarly publications
Rank 3: Student-generated content	Rank 3: Archival content
Rank 4: Digitized images	Rank 4: Electronic theses and dissertations
Rank 5: Electronic theses and dissertations	Rank 5: Digitized images

CARLI institutions have opinions about the features that should be available in a consortial IR. (See Appendix I, Table 8.) The top ten desired features are, in order from most desired to least:

1. Usage statistics
2. Customizable metadata
3. Unlimited storage
4. Institutional branding
5. Display media formats (audio, visual)
6. Integration into campus authorization/authentication systems
7. Suppress content from public view (i.e., campus-only viewing)
8. Search engine optimization (SEO)
9. Batch upload
10. Electronic Theses and Dissertations (ETD) workflow

Migration concerns

There are several different dimensions of migration that warrant consideration after a platform has been selected: communication and outreach, technical capacity, metadata concerns, post-migration staffing and support, setting guidelines for participating institutions, and managing expectations overall.

CARLI may want to consider and plan for supporting the communication and outreach needs of institutions choosing to migrate to a shared platform *before* such a migration would take place. A cohesive marketing campaign to faculty and administrators statewide, e.g., a CARLI site with an FAQ, brochures, bookmarks, etc. could be an opportunity for CARLI to underscore its importance to providing worldwide access to scholarship by Illinois faculty.

With regards to platform, in reviewing statewide data, the preponderance of existing IRs utilize two systems: DSpace or bepress. Migration concerns would be reduced if one of these two systems were chosen for a shared effort, but if a third system were chosen, this would mean migrating data from numerous instances of both bepress and DSpace to the chosen CARLI shared IR platform. Given these considerations, it is worthwhile at this exploratory stage to investigate migration challenges native to both the DSpace and bepress platforms.

Fortunately, the DSpace and bepress data models for repository objects are both rather simple. At their core, a repository object in either system consists of a deposited file or files and their associated descriptive and administrative metadata. DSpace and bepress also feature their own utilities for exporting these digital objects and their metadata, making them on the whole easily amenable to migration to other platforms.

The challenge lies in the way these digital objects are enmeshed in a number of ancillary features sure to complicate migration to another system. These include:

- 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.

Most importantly, while the data models for the DSpace and bepress systems are rather simple, anyone who has led a migration of content from one system to another knows that the process is inevitably fraught with unexpected problems. Many institutions that run either DSpace or bepress have introduced their own local oddities, ranging from different uses of metadata profiles to more invasive customizations. These all need to be taken into consideration during a migration and involve both intellectual labor from those who understand local practices of member institutions as well as technical labor from whomever is managing ingest into a new system. Realistically, if a large-scale migration of IR content were to occur across the CARLI membership, we would

require a metadata analyst and a technical staff member proficient in data transformation and file packaging to dedicate considerable effort to such a process.

There are also post-migration matters to consider in terms of staffing and supporting CARLI institutions. In one case at Florida State University (FSU), the decision to migrate from bepress was due to the increasing costs of a hosted solution and a desire to support open source platforms[4]. Islandora was already supported by the Florida Academic Library Service Cooperative (FALSC), which represents college and university libraries within the Florida Virtual Campus (FLVC) system, primarily to support digital collections and special collections. FSU's Special Collections was already using Islandora for their digital collections, so there was familiarity with the platform, internal expertise, and additional support from the state to develop Islandora to fit the needs of a repository. While there was some concern about faculty reaction to the migration, there have been no complaints since the new repository went live. FSU passed an Open Access Policy in February 2016, so submissions from faculty greatly increased post-migration, which helped to identify issues with submission and administration.

FSU Libraries employs a Linux system administrator as well as two developers, but their time is dedicated to FSU library projects. FALSC provides most of the support for Islandora across the state for all eleven Islandora sites, employing one Linux system administrator and one software developer who work on other projects as well.

Islandora does not provide all the features and functionality available in bepress. User submission and administration on the backend are not as user friendly, and visualizations such as the real time readership map and the dashboard are at present not a feature in Islandora. At this time, there is no module for journal publishing; FSU migrated their journals from bepress to a consortially hosted OJS instance at around the same time as the IR migration. However, there are modules for specific types of content (e.g., audio, video, compound objects, etc.) as well as for usage statistics. Events and author profiles can be custom built with developer time. SEO is similar to bepress. The FALSC Islandora users' wiki is a rich source of information for every aspect of the system.

With the Elsevier acquisition of bepress, there has been a significant interest within the Islandora community to develop a more robust institutional repository solution. Efforts are underway to identify features and functionality from the Islandora community and developer time has been 'donated' towards this effort.

[4] http://purl.flvc.org/fsu/fd/FSU_libsubv1_scholarship_submission_1462290278

Analysis

The landscape of IR solutions is in a period of transition. Key factors include:

- **Uncertainty around bepress Digital Commons**. Elsevier's 2017 acquisition of bepress has elicited a certain amount of concern in libraries[5], and it remains to be seen whether Elsevier will change bepress's Digital Commons product and licensing fees. While a number of bepress users wish to remain with the product, many are seeking alternatives.
- **Uncertainty around the future of DSpace**. Frustration has been mounting for years within the DSpace community with many of the software's limitations. It remains to be seen whether DSpace will respond with dramatic changes to the software itself, or if an exodus of DSpace users will see a surge in the quality of competing platforms.
- **The immaturity of Fedora-driven solutions**. There is great potential for the future of systems based on the Islandora or Hyku stacks, but it is broadly acknowledged by those who are using them that these are far from mature solutions for a consortial IR.
- **The emergence of additional options.** With people moving away from bepress and DSpace, and many finding Fedora solutions unwieldy, additional options may emerge in coming years from the academic and open source software community to fill the need for shared repository services. In particular, the University of Illinois at Urbana-Champaign is moving away from the DSpace system currently powering its IDEALS (https://www.ideals.illinois.edu/) IR, and is embarking on building a new solution on top of its existing "Medusa" repository platform for preservation and access. Its repository team and library director have indicated an interest in building this system in a way to allow its use by CARLI membership, and have stated a willingness to play a leadership role in deploying a statewide IR solution (for more information, see Appendix IV). They estimate two to three years as being necessary to complete much of the required technical work and planning.
- **The push into cloud-driven architectures**. There is a general trend in the digital library community to move away from locally managed server infrastructure in favor of a cloud driven model, e.g. hosting repository services in Amazon Web Services. Leaders in the field of digital libraries are still working out the kinks of this shift, with the expectation that future solutions will be more easily deployable and adoptable by institutions or consortia that do not care to manage server technology locally.

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 than it is right now.

[5] See "beprexit," https://beprexit.wordpress.com/

Recommendation

The IRTF has determined that there is compelling interest throughout Illinois in a consortial IR solution, but that the international market for IR solutions, both open source and vendor-based, is in flux. The IRTF is confident; however, that the landscape of IR technology will stabilize in the coming two to three years, presenting a much clearer slate of options than at present. For this reason, the IRTF recommends that CARLI keep the door open to a statewide IR solution by retaining some version of this Task Force with the goal of periodically monitoring the technical landscape and facilitating relationships among CARLI members with those member institutions, such as the University of Illinois at Urbana-Champaign, that have expressed interest in assuming a leadership role in developing consortial solutions.

For a more in-depth view of the IRTF's findings, please see its responses below to its initial list of possible recommendations.

1. CARLI does or does not act as an IR provider for CARLI members at this time.
   Our research demonstrated that there is significant interest in pursuing a consortial institutional repository among CARLI members. Unfortunately, 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. Our group recommends that CARLI does not act as an IR provider for CARLI members at this time. We recommend that CARLI continue to monitor the IR space to see how the market evolves. Our investigation of vendors indicates an interest in pursuing consortial IR options, but these are only just beginning. CARLI could actively pursue those conversations with vendors, but the Task Force does not recommend it at this time. 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. 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.

2. CARLI acts as a clearinghouse for information and education about IRs but not provide IR functionality for CARLI members.
   The CARLI office and staff is currently unable to provide this service. Additionally, we feel that, while certain elements of service provided by such a clearinghouse, the educational component in particular, overlap with the efforts of the CARLI Collection Management, Created Content, and Public Services Committees, this work would not fit into a single committee Our group recommends retaining some form of the current Task Force to monitor the evolving landscape of institutional repository platforms and to generate limited outreach in the form of updates or an email list for members on this topic.

3. CARLI does not act as an IR provider, but helps to facilitate (e.g., provide meeting space, an email list, announcements) member libraries to pursue shared IR opportunities.
   Our group recommends that the revised charge to the Institutional Repositories Investigative Task Force include language about acting as a facilitator to member libraries to pursue shared IR opportunities, in addition to the informational and educational components outlined above.

4. CARLI acts as the procurement agent for an IR and offers the product to the members similar to brokered electronic resources.

   Our investigation unearthed no evidence that vendors are interested in pursuing these types of relationships, and there is no indication that this option would be more cost effective for CARLI member institutions. These reasons, paired with CARLI's prior experiences negotiating brokered platforms, leads us to believe that this could be a very time-consuming endeavor with limited benefit. The Task Force could continue to monitor this option, but at this time, our group does not recommend that CARLI acts as the procurement agent for an IR.

5. Specifications in a Memorandum of Understanding to which a group could respond, in which CARLI would act as a broker.

   Our group did not find an institution in the state that is willing or able to host another member's IR at this time. As part of the next phase of the Task Force, that group could continue to monitor this possibility and, if necessary, provide templated language for inclusion in a Memorandum of Understanding that CARLI could broker. Our recommendation is to not pursue this option at present in anticipation of a scenario arising, but rather to wait for a scenario to arise and provide expertise at that point.

Appendix I: Full Survey Results

CARLI Institutional Repository Investigative Task Force Survey Results

The Task Force surveyed CARLI members from January 8, 2018 through February 2, 2018 to find out:

- Use of and satisfaction with institutional repositories (IRs)
- Reasons an IR has not been implemented
- Interest in and anticipated use of a CARLI consortial IR

The survey went to directors of the 134 CARLI members; 74 institutions responded with usable answers to all or part of the survey, for a response rate of 55.2%. The institutions who responded to the survey were coded with a Carnegie Classification:

Table 1

Institution type	Institution type represented in CARLI	Survey respondents
Baccalaureate	21 (15.7%)	9 (12.2%)
Community College	39 (29.1%)	19 (25.7%)
Doctoral	13 (9.7%)	10 (13.5%)
Masters	26 (19.4%)	18 (24.3%)
Special	35 (26.1%)	18 (24.3%)

Twenty-eight respondents (38%) indicated their institutions currently have an IR and 45 institutions (62%) do not have an IR (n=73).

Current use of and satisfaction with institutional repositories

Of the 28 respondents that said their institution has an IR, 27 provided information on the platform and their level of satisfaction. 33.3% of respondents were “very satisfied;” 51.8% of respondents were “somewhat satisfied;” and 14.8% were “not satisfied.”

Table 2

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

Comments:

- This is a shared, consortia-based IR, and as such we are limited in the decision-making processes for individual changes.
- Would prefer to have a little more local control, but overall satisfied.
- We were happy with our relationship with bepress until the sale to Elsevier. Now we have serious concerns about it continuing to exist as a true open access repository.
- Since this is a shared IR with other LIBRAS institutions and administered by the IT department of one of the institutions we have had some difficulty with response time for problems with DSpace.
- The college archives needs a server. We are having a difficult time convincing the College I.T. dept. We recently remodeled, and could not rope the cost of a server into the remodel. I am not very hopeful. We need at least 2 TB of space for video. My biggest concern is the "born digital" video that we cannot capture, such as our president's all campus addresses and graduations. These are stored on livestream. [identifying URL removed] That is just one example. Speech and theatre and Sports/Wellness have their recordings house in different places. The Archives is unable to capture these videos.
- Currently very satisfied, however, concerned about potential changes to bepress policies, pricing structures, etc. as a result of sale to Elsevier.
- Would like to have more flexibility with metadata schema and discipline taxonomy, and ability to stream files (I believe bepress is working on streaming in 2018).
- We are currently migrating to a new platform, Islandora.

- We struggle with the demand for supporting large data deposits and delivery, as well as custom metadata requirements for student projects (e.g. department names, adviser names, grant awards) and esoteric research (e.g. audio organized by speaker, age, language, dialect).
- We launched the IR with the intent of hiring a tenure-track librarian to manage it, taking over for the launch team. Then, the hiring freeze hit.
- Our project implementation has really stalled due to personnel issues...we just don't have the IT support and the library person spearheading this left to take another job.
- Our IR solution fulfills the minimum requirements of an institutional repository. We have been unable to connect our repository to Google Analytics. Also, the repository is not responsive to different browser sizes.
- Subject indexing is not fully implemented.
- Support is generally very good and reasonably fast. Would prefer if they weren't an arm of Elsevier, of course.

Current IR content

Respondents were asked about the amount of data in their IRs.
Twenty-one of the 28 institutions with an IR responded with usable information about estimated number of files. This number ranged from a low of 40 files to a high of 78,973 files. Total number of files: 167,703.

Fewer respondents—only 10—were able to estimate the amount of data in GB in their IRs: a low of 1 GB to a high of 550 GB. For those 10 respondents, the total amount of data indicated is 1034 GB.

Respondents were asked to select the type of content currently held in their IRs (n=28):

Table 3

Type of content	No. of respondents	% of respondents
Student-generated content	22	78.6
Scholarly publications	21	75.0
Archival content	18	64.3
Electronic Theses and Dissertations	18	64.3
Digitized images	14	50.0
Administrative documents	13	46.4
Audio	13	46.4
Conference and events contents	13	46.4
Journals	13	46.4
Video	13	46.4
Research data	11	39.3
Teaching materials	9	32.1
Non-institutional community-generated content	4	14.3
Other*	5	17.9

*Comments for “Other” include:

- Campus governance documents
- We have audio and video housed in third-party systems (e.g., YouTube, SoundCloud, Vimeo), and embedded in some collections in our IR. Copies of the videos are also kept on a library server.
- Event management
- Faculty-written scholarship
- Exhibits, artwork

Reasons an IR has not been implemented

Of the 45 respondents that said their institution does not have an IR, 42 provided information on reasons for not having an IR. Below is the list of selections and the count of the respondents that selected each reason. Respondents were able to select more than one reason.

Table 4

Reason	Count
Lack of funding to support ongoing storage and maintenance	29
Lack of staff support	22
Among competing priorities, IR implementation has not risen to the top	21
Lack of technical expertise	17
Other	11
We tried to implement an IR and failed	1

Below are the comments grouped roughly by reason for not having an IR. There were also comments specific to existing platforms and services that are broken out at the end of this section.

Resources / support

- We simply don't have the resources to implement IR
- We are such a small institution, this is beyond our resources.
- We just don't have the financial or staffing resources to create and maintain an IR.
- Lack of funding for up-front costs.
- There is no money for this.
- We have very limited staff resources and are already stretched thin managing the various systems and services we already have. In addition, although we have been very interested in the potential of an IR for a number of years, there has been uncertainty about adding yet another system/service with limited staff, coupled with lack of a compelling demand, e.g. from faculty, to support such a system/service.
- Due to lack of funding and lack of staff time to support an IR, we have not yet begun to investigate how to establish or maintain a digital repository.
- Lack of administrative support. We have lots of buy-in from faculty and deans, but not the provost's office.
- We would like to get an IR going, but have no commitment as to either staff or budgetary support at this time.

Priorities

- It's a “nice to have,” but as a community college, faculty research is not a priority.
- I have not approached an IR in a serious manner with upper administration, but I recognize the need for one.
- This is on the radar for us, but I think we might pursue software for building faculty profiles first, and hopefully end up with sort-of an institutional repository "index" as a result. Having an accurate publication list would be a good first step. Then we can explore proper repository activities (we have a lot of research activity here, so it would be nice). We'll likely need some funding/support (from the institution--not CARLI) to get to that point, though.
- No need at our institution.
- We have not seen a need for one at this time.
- No stated interest from our institution's seven libraries or other departments. Also, it would require a commitment in terms of funding and staff that's probably not available in our current budget environment.
- Given staffing and budget reductions, we haven't been able to make an IR a priority.
- We don't want to digitize our institutional documents.
- An IR was created six years ago with the help of grant funding but was not maintained beyond the initial grant period because of a lack of funding. At this point, we are so short-staffed that creating an IR has fallen to the bottom of our list.
- I am not aware of any conversation at any level about an IR.

In Process

- I put into the budget an IR subscription to begin in July 2018. This budget has not yet been approved.

Digital Commons

- We wanted to use Digital Commons, but it's not within the library budget.
- We had looked some at Bepress, but it was more expensive than we could afford.

CONTENTdm

- We are using CONTENTdm for some institutional collections. We have not yet discussed a stand-alone IR for additional materials/purposes

LIBRAS

- We are part of LIBRAS, and LIBRAS has a shared institutional repository which we were planning to use. However, technical issues have prevented us from having access.
- We are a member of LIBRAS Consortium and have access to the Constellation Digital Repository but have yet had the necessary time or ability to devote to its implementation.
- We do have access to the LIBRAS IR now as Constellation but have not made use of it.

Interest in and anticipated use of a CARLI consortial IR

The survey asked “If CARLI were to offer a consortial IR solution, would you consider a cost-sharing arrangement with CARLI for your organization to benefit from access to a shared platform and training opportunities?” (n=66). The institutions who responded to the question were coded with a Carnegie Classification. *Note: This question was available both to those respondents who currently have an IR and to those who do not.*

Table 5

	Yes	No	I don’t know
Baccalaureate	7	1	1
Community college	8	2	6
Doctoral	3	3	3
Masters	12	0	5
Special	8	2	5
Total	**38 (57.6%)**	**8 (12.1%)**	**20 (30.3%)**

Respondents were asked to prioritize functions of a consortial IR (n=38). Respondents ranked the eight functions against each other on a scale of 1 to 8, with 1 being the highest priority and 8 the lowest priority.

Table 6

Function	Average ranking
Provide access to student work	2.6
Provide access to faculty research publications	3.3
Provide access to digitized collections	3.5
Provide digital preservation service for institutional content	4.3
Provide access to faculty research data	4.5
Provide access to campus media content	5.4
Showcase bibliographies of faculty work	5.5
Comply with government mandates for open access to research	5.7

Respondents were asked to select the type of content would be housed in the consortial IR (n=35). Respondents were able to select one than one content type.

Table 7

Content	Count of responses
Archival content	29
Scholarly publications (e.g. journal articles, book chapters)	24
Student-generated content	23
Digitized images	21
Electronic Theses and Dissertations	20
Administrative documents	16
Teaching materials and learning objects (textbooks, syllabi, open educational resources)	16
Audio	13
Video	12
Research data	11
Conference and events contents	10
Journals	10
Other*	4
Non-institutional community-generated content	2

*Comments for “Other” were versions of “Unknown” or “None.”

Respondents were asked to provide an estimate of the number of files they currently have ready for deposit. Only 21 of the 38 respondents answered this question. Twelve of those responses were “Unknown” and four were “None.” Of the remaining responses the answers ranged from 10 to 2000 files.

Respondents were also asked to provide an estimate of the amount of content they currently have ready for deposit. Only 19 of the 38 respondents answered this question. Twelve of those responses were “Unknown” and four were “None.” Of the remaining responses the answers ranged from 1 to 500GB.

Respondents were asked to prioritize the features that they would like to see in a consortial IR. The respondents were asked to assign a priority to each feature listed. The responses ranged from “No opinion” to “Low,” “Medium,” or “High priority.” For purposes of analysis we converted the responses to the following values:

- No opinion = 1
- Low priority = 2
- Medium priority = 3
- High priority = 4

And then calculated the sum, average, and median (converted to response) below.

The high priority features are listed below:

Table 8

Feature	Sum	Average	High priority	Low priority
Usage statistics	146	3.84	31	0
Customizable metadata	136	3.58	22	1
Institutional branding	125	3.38	20	0
Integration into campus authorization/authentication systems	126	3.32	18	0
Electronic Theses and Dissertations (ETD) workflow	116	3.05	18	10
Unlimited storage	127	3.43	17	1
Display media formats (audio, visual)	124	3.35	16	3
Suppress content from public view (i.e., campus-only viewing)	118	3.11	16	10
Search engine optimization (SEO)	118	3.11	15	5
Batch upload	118	3.11	13	5

The remainder of the features (9) were listed as medium to low priority and we have added a count of respondents that listed this feature as a high priority and a low priority to the table below

Table 9

Feature	Sum	Average	High priority	Low priority
Visualizations of users (readership map, download maps)	109	2.87	9	10
Embargo content	101	2.73	10	12
Journal management and production tools	103	2.71	9	10
Self-deposit	107	2.82	9	14
Integration with third-party software (e.g., Digital Measures, Vimeo, YouTube, etc.)	104	2.81	7	10
Research data management tools	98	2.72	7	11
Integration with faculty profile systems	90	2.37	6	5
Tools for managing conferences or events	87	2.29	1	1
Is based on open source software	86	2.26	3	9

Respondents were also allowed to list other features and leave comments.

- Display of use/permissions/copyright policy, when applicable.
- Very user-friendly public interface.
- Long term commitment to open access and a consistent support on the backside since staffing at individuals libraries will be an ongoing issue.

Finally, respondents were asked to prioritize the services that they would like to see in a consortial IR. The respondents were asked to assign a priority to each feature listed. The responses ranged from “No opinion” to “Low,” “Medium,” or “High priority,” and the responses were analyzed in a manner similar to the features. Unsurprisingly, all services had a median value of high priority.

Table 10

Service	Sum	Average
Institutional control over content organization and management	143	3.76
Documentation of system and how to use it	139	3.66
Readily available technical support	139	3.66
Training	138	3.63
Migration support	126	3.32

Respondents were also allowed to list other services and leave comments.

- While CARLI covering support, training, and documentation services might be nice, surely the IR provider could supply this, so CARLI would not need to...or possibly a user group could provide some this via web pages, workshops, or email lists.
- CARLI always does a fabulous job.

Lack of interest in a consortial IR service

Respondents who responded “no” to the question, “Would you be interested in a consortial IR?” were asked why they would **not** consider a consortial IR. The question was an open-ended dialog box (n=25), but some respondents cited more than one reason. The most common reasons for not considering a consortial IR:

Table 11

Reason	No. of respondents citing
Cost	10
Already have an IR	8
No need/interest	5

Some comments:

- We already have a consortial solution (LIBRAS). Plus the last solution CARLI provided that was similar to a IR in scope (SharedSpaces) was unwieldy and complex.
- Have one already -- it would need to offer benefits not already receiving from the current IR we are using as well as to be able to transfer existing files into the new system.
- It is a very low priority, our library and university are struggling to maintain basic services, extended services are just not practical.
- We have an established relationship with bepress with customizations that may be difficult to transfer/emulate in a consortial IR solution. Also, we share costs with three other offices on campus, which has been an advantage for us not only financially, but also in communicating the value of the IR across campus. Barring any unforeseen financial exigency in the future, we're happy with our IR solution at this point.
- We are in the process of migrating repository platforms and have put in work to analyze and implement our institution's specific needs for the new repository.
- Our present solution works well -- only interested if CARLI offers discount on our existing platform.
- We use an open source IR codebase (Hyrax) and our development on the codebase is already constrained by the needs of the open source community (Samvera). It would make things more complicated for us if our work and priorities had to be reconciled with an additional community's needs and priorities.
- Because of our temporary stall, we don't have a sense of how important university branding is to our faculty who will contribute.
- Security, cost, duplication of effort
- Depends on cost and on level of support. Current support is very good.
- Much would depend on cost and whether or not the IR solution would include a mechanism for the long-term preservation of the content.

General comments

Additional general comments were provided by some respondents:

- IRs are used for so many purposes (some of which are very similar) that it would be hard to envision a one-size-fits-all solution. Also, since IRs are used by many institutions for site archival purposes, we would be entering into a potential very long-term contract.
- Overall, I think it's a good idea and would benefit CARLI institutions who do not have a solution presently and might be useful in the ongoing and upcoming OER initiatives being considered by CARLI.
- This is a great idea generally. This is exactly the sort of things that CARLI should be involving itself with!
- Our institution is also aware of, and supports, the [name of IR removed]. This is a digital library for our denomination and several institutions also have instances of it as their own IRs, although we still have our own IR in Digital Commons.
- I'm glad this is being explored. If we can keep the functionality of our bepress platform and have the long term commitment to open access, we would be thrilled.
- How is this different than CARLI Digital Collections? Would that go away? It's a little clunky, in my opinion, or we are not using it correctly.
- While we are willing to share cost, the cost would need to be low for us to be able to participate!
- If a grant was available for staffing and IT help, that might allow [institution name removed] to participate.
- Thanks for looking into this!
- It would be great if CARLI would swing a deal with Digital Commons to control price increases, and to make some sort of consortial community
- I would caution against community-based solutions that attempt to serve the needs of digitized collections (i.e. images, books) and born-digital research materials (working papers, data sets). Both require specific metadata, organizational, and display considerations and the decision to accommodate both sets of use-cases could lead to slow deliberative processes and limit an institution's ability to respond to campus needs.
- Faculty, directors and deans have recognized this as a huge need on campus. We have been working on this for two years but cannot secure funding for our own platform due to lack of institutional priority at the top level. A more affordable consortial solution would be a major win.
- [Institution name removed] currently maintains an IR and would be interested in working with CARLI for cost savings and efficiency.
- Visually pleasing and user-friendly interface.
- We intend to grow the IR to include masters theses, events and conferences, and journal publishing.
- Would need a strong User Group Advisory Council.

- We're glad you're thinking about this!
- Discussions beyond the campus library have pointed to a need for a robust document management system. The library and an IR would be a viable solution with the right support (and partnering with our campus IT).

Topics for further investigation

The Task Force discussed the results of the survey and determined there are areas to investigate further such as the intentions and commitment of the respondents. The survey demonstrated interest in a consortial IR solution for CARLI membership with 38 institutions responding “Yes” to the question, “If CARLI were to offer a consortial IR solution, would you consider a cost-sharing arrangement with CARLI for your organization to benefit from access to a shared platform and training opportunities?,” but the survey was not able to gauge the degree of that interest or the level of cost share that the institutions would be able to devote for this initiative. With that in mind, the Task Force will continue our investigation and present a final report to the CARLI Board in June.

Appendix II: CARLI Institutions with IRs

Institution	Platform	IR URL
Eastern Illinois University	Digital Commons	http://thekeep.eiu.edu/
Southern Illinois University Carbondale	Digital Commons	http://opensiuc.lib.siu.edu/
Illinois Wesleyan University	Digital Commons	http://digitalcommons.iwu.edu/
DePaul University	Digital Commons	http://via.library.depaul.edu/
Governors State University	Digital Commons	http://opus.govst.edu/
Illinois State University	Digital Commons	http://ir.library.illinoisstate.edu/
Columbia College Chicago	Digital Commons	http://digitalcommons.colum.edu/
Parkland College	Digital Commons	https://spark.parkland.edu
Lake Forest College	Digital Commons	http://publications.lakeforest.edu/
University of Chicago	Digital Commons	http://chicagounbound.uchicago.edu/
Loyola University Chicago	Digital Commons	http://ecommons.luc.edu/
College of DuPage	Digital Commons	http://dc.cod.edu/
Illinois Math and Science Academy	Digital Commons	http://digitalcommons.imsa.edu/
Augustana College	Digital Commons	http://digitalcommons.augustana.edu/
Southern Illinois University Edwardsville	Digital Commons	http://spark.siue.edu/
Olivet Nazarene University	Digital Commons	http://digitalcommons.olivet.edu/
National-Louis University	Digital Commons	http://digitalcommons.nl.edu/
Northeastern Illinois University	Digital Commons	http://neiudc.neiu.edu/
Northwestern University	Hyrax	https://arch.library.northwestern.edu

	(Samvera)	
Benedictine University	DSpace	http://ben.constellation.libras.org/
Dominican University	DSpace	http://dom.constellation.libras.org/
Elmhurst College	DSpace	http://elmhurst.constellation.libras.org/
Harper College	Homebrew solution	http://dept.harpercollege.edu/library/archives/repositoryhome.html
Illinois Institute of Technology	DSpace	http://repository.iit.edu/
John Marshall Law School	Digital Commons	https://repository.jmls.edu/
Judson University	DSpace	http://judsonu.constellation.libras.org/
Lewis and Clark Community College	Digital Commons	http://www.lc.edu/Digital_Commons/
Lewis University	DSpace	http://lewisu.constellation.libras.org/
North Central College	DSpace	http://northcentral.constellation.libras.org/
Northern Illinois University	DSpace	https://commons.lib.niu.edu/
Roosevelt University	CONTENTdm	http://collections.carli.illinois.edu/cdm/landingpage/collection/rou_urr
Saint Xavier University	DSpace	http://sxu.constellation.libras.org/
Southern Illinois University School of Medicine	Digital Commons	http://opensiuc.lib.siu.edu/
University of Illinois at Chicago	DSpace	https://indigo.uic.edu/
University of Illinois at Springfield	DSpace	https://www.ideals.illinois.edu/handle/2142/13852
University of Illinois at Urbana-Champaign	DSpace	https://www.ideals.illinois.edu/

Appendix III: Meeting Summaries

November 11, 2017 - teleconference

- Reviewed and discussed of the Task Force charge.
- The Task Force decided to create one survey for the membership to gather information related to Task Force Charges #2 (Conduct an environmental scan) and #3 (Assess the existing IRs platforms, inventories, and institutional guidelines) within the CARLI membership members' repository structure and inventories.

December 6, 2017 - CARLI Office

- Discussed the scope of IR definition.
- Created a draft document of the Member Survey in Google Docs and began drafting questions for the survey.
- Established deadline for Task Force to complete survey.
- Established dates when survey would be sent to libraries and response deadline.
- Discussed future steps and meeting dates.

January 27, 2018 - teleconference

- Discussed initial survey responses.
- Updates and discussed the progress of work related to the Task Force Charge, including the environmental scan, software platforms and services, migration issues, feasibility and structure of a consortial IR.

February 14, 2018 - Parkland College

- Discussed survey results.
- Discussion regarding bepress services.
- Created a preliminary report draft.

February 27, 2018 - teleconference

- Reviewed report draft.
- Discussed examples of existing consortial IRs: Texas Digital Library, California State, OhioLink, migration issues and shortcomings.
- Discussed Task Force recommendations for CARLI.
- Group decided to delay final report until June Board meeting, submitting preliminary report instead.

April 10, 2018 - teleconference

- Received feedback from Board about preliminary report and addressed questions.
- Discussed possibility of conducting a follow-up survey to respondents who currently have IRs to gather more data on IR expenses. Opted for informal email for data collection.

May 11, 2018 - teleconference

- Discussed responses to follow-up email to survey respondents.
- Identified responses to each of the six possible recommendations.
- Assigned tasks to complete final report.

Appendix IV: IR Development at Illinois

The University of Illinois at Urbana-Champaign (UIUC) recently committed to migrating its institutional repository (IDEALS, https://www.ideals.illinois.edu/) from its underlying DSpace software to its own locally developed “Medusa” stack of preservation and access services. UIUC intends to build this next-generation IR platform in a cloud environment powered by Amazon Web services, with an eye to opening them in some form to external institutions. As the state’s flagship land grant institution, UIUC proposes to share its system and service management expertise with other institutions in the state, and to develop its next-generation IR solution as an option for consideration to the CARLI membership in the coming two to three years.

The approach so far has been to design durable repository architecture with distinct curation modes and services built on top of it, coupling stable infrastructure for the storage and management of digital objects with a small but carefully chosen set of tools for describing and providing access to those objects. Accomplishments have included:

- Provisioning the robust Medusa (https://medusa.library.illinois.edu/) digital preservation service for locally managed digital collections, securing control of over 110TB (10,000,000 files) of content.
- Placing Medusa at the heart of local repository architecture that makes digital preservation an essential element of ongoing workflows, rather than an isolated, unconnected silo.
- Building curation and access services like the Illinois Data Bank (https://databank.illinois.edu/) and the Digital Collections (https://digital.library.illinois.edu/collections) on top of Medusa architecture to allow for a variety of data models, metadata profiles, and presentation of digital content.

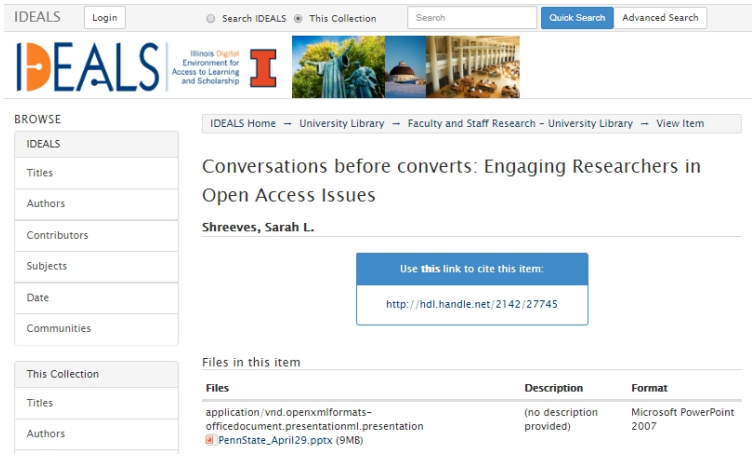
These efforts have taken place in two phases to date, under the banner of the Medusa project:

- Phase I (2012-2015): build Medusa collection registry, gain basic control over the library’s locally managed digital assets by describing them in the registry and placing the files in preservation storage.
- Phase II (2015-2018): bring greater structure to collections, build access mechanisms on top of Medusa storage such as the Digital Library (https://digital.library.illinois.edu/) and Illinois Data Bank (https://databank.illinois.edu/).
- Phase III (2018-2021): migrate the IDEALS institutional repository to the Medusa stack of services with multi-institutional capacity built-in, and to establish these services as a cloud-driven Amazon web service with an eye to potential use by institutions outside the University of Illinois.

More information on UIUC repository systems and services is available below.

IDEALS

IDEALS, the Illinois Digital Environment for Access to Learning and Scholarship, collects, disseminates, and provides persistent and reliable access to the research and scholarship of faculty, staff, and students at the University of Illinois at Urbana-Champaign.

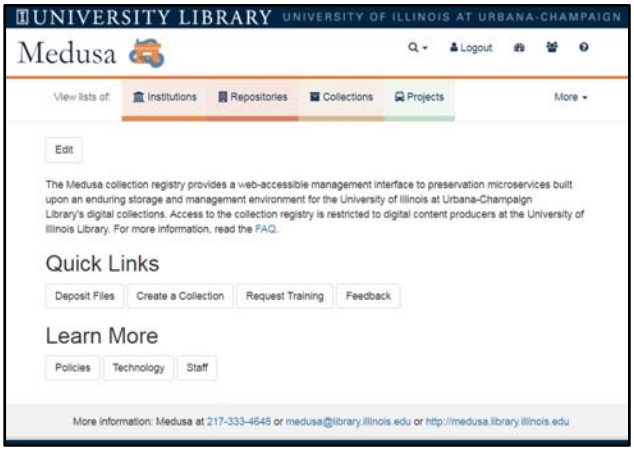


Faculty, staff, and graduate students can deposit their research and scholarship—unpublished and, in many cases, published—directly into IDEALS. Departments can use IDEALS to distribute their working papers, technical reports, or other research material.

IDEALS is available at https://www.ideals.illinois.edu/ and also includes all electronic theses and dissertations (https://www.ideals.illinois.edu/handle/2142/5131) deposited from Fall 2010 onward.

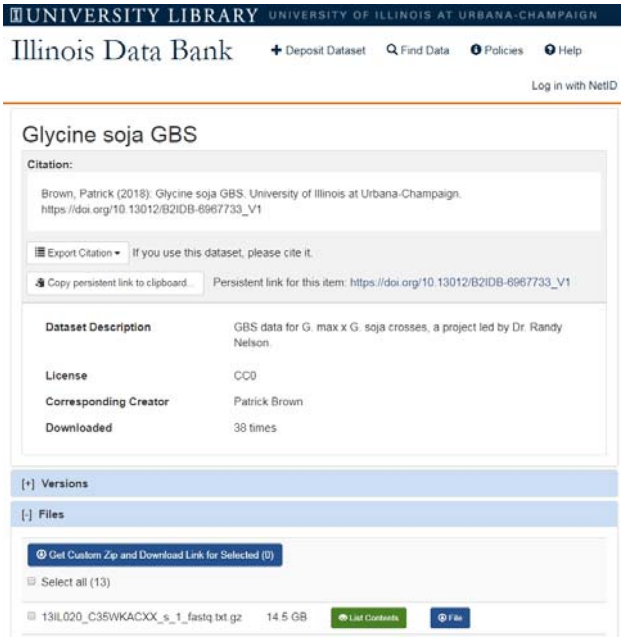
Medusa

The Medusa digital preservation repository provides an enduring storage and management environment for digital collections. At present, Medusa's collecting focus is on digitized and "born digital" books, manuscripts, photographs, audiovisual materials, scholarly publications, and research data from the library's special collections, general collections, and institutional repository. All master files created by the library's digitization units, for example, are by default deposited into Medusa. Medusa's storage infrastructure consists of two copies of every file, replicated across two distinct campus nodes and a third copy of every file backed up off-campus, although this is subject to change as the library moves Medusa into Amazon Web Services. While most of Medusa is off-limits to the public, its live production site is available at https://medusa.library.illinois.edu/.



Illinois Data Bank

The Illinois Data Bank's mission is to centralize, preserve, and provide persistent and reliable access to the research data created by affiliates of the University of Illinois at Urbana-Champaign, such as its faculty, academic staff, and graduate students. The Research Data Service at the University Library maintains and operates the Illinois Data Bank on behalf of the University of Illinois at Urbana-Champaign campus. The University Library has a long history of successfully supporting the preservation of and access to Illinois research outputs. The Illinois Data Bank operates within a robust policy framework that fully describes the University's commitment to providing persistent and reliable access to research data. It is available at https://databank.illinois.edu/.



Digital Collections

The digital collections of the University of Illinois at Urbana-Champaign Library provide web access to many of the Library's signature holdings for students, scholars, and the general public. They include digitized books, photographs, newspapers, maps, and other research materials.

The digital collections site features collections powered by a Digital Library System built directly on top of Medusa services, making for seamless integration between digital preservation and access services, as well as links out to collections that live in external systems. This clearinghouse of our locally managed digital collections is available at https://digital.library.illinois.edu/collections.