

# Institutional Repositories: An Analysis of Trends and a Proposed Collaborative Future

by Leila Sterman

This is the Author's manuscript (Postprint) version of an article published in College & Undergraduate Libraries on 11 Sept 2014.

Sterman, Leila. (2014) "Institutional Repositories: An Analysis of Trends and a Proposed Collaborative Future" College & Undergraduate Libraries  
Vol. 21, Iss. 3-4, 2014 <http://doi.dx.org/10.1080/10691316.2014.943919>

## Abstract

*This study seeks to give libraries a plan for inter-institutional cooperation for institutional repositories that will benefit all involved: researchers, institutions, and, ultimately, global scholarship. This research uses repository studies, interviews with existing repository managers, and the input of libraries considering a repository to inform the exploration of the opportunities for collaboration in IR development and maintenance. This paper proposes opportunities for collaboration between institutions in order to convince libraries that it is possible and effective to work together toward a common goal: highlighting existing working groups or alliances, sharing technology and hardware, building separate inter-institutional bodies to house repositories, and sharing the work of specialists.*

Key Words: Open Access, Digital Repository, Collaboration

## **Introduction**

This paper seeks to help libraries join together to form more efficient and useful collaborations for the sustainable success of institutional repositories (IRs). Collaborative efforts can help form more efficient and systematic management of IR services.

Repository managers may then have more time and resources to put toward solving a number of problems or challenges. These include: increasing indexing rates by search services such as Google Scholar (Arlitsch and O'Brien 2012); integrating unique identifiers like Open Researcher and Contributor ID (ORCID) and the International Standard Name Identifier (ISNI) system; engaging in outreach to increase support, understanding and funding for IRs; and soliciting scholarship to populate the IR.

## **The State of Repositories**

### **Background**

As the budgets shrink in many libraries, there has been a shift toward e-resources in an attempt to focus resources where they will provide the greatest return on investment within institutions (Oder 2009). In a paper describing the institutional repository at Pacific University, the authors note that “in taking on a much more active role in the creation, dissemination and preservation of internally produced scholarship, the Library has demonstrated its value to faculty and administrators and has opened the door to new partnerships which will not only strengthen the University, but also the Library’s place within it” (Gilman, Gilman, and Kunkel 2010, 22). Gilman, Gilman, and Kunkel further suggest that the academic library move from being an access point and archive of information to a place to increase engagement in the scholarly communication life cycle. Bankier and Smith speak to the wave of new technology-supported initiatives that are possible in an academic library. They state; “by seeking out a variety of content types, the library is able to initiate, renew, or redefine its relationship with faculty, departments, and administration, generating critical support for scholarly communication and repository initiatives” (Bankier and Smith 2010, 245). Librarians have a responsibility to be aware of and advocate for new technology-- especially technology such as IRs that have the ability to bring such a great deal of information to such a broad range of people.

Although an institutional repository could become a useful and important tool for universities, not every institution has the means to create and maintain their own repository. Repositories are expensive and time consuming, and they demand specific knowledge of programming, content management, metadata applications, publicity, and internal marketing to researchers.

In the past few years, IRs have proliferated within the context of the open access(OA) movement, shrinking library budgets, the increasing costs of traditional scholarly journals and publications, and the increased development of software platforms for IRs. In 2014, The Ranking Web of World Repositories lists 272

repositories in the USA, and 1746 repositories in the world (Cybermetrics Lab 2014). Since arXiv, one of the first repositories and the top ranked repository in the Ranking Web, began in August 1991, there has been a relative explosion of digital repositories in the academic world (Cornell University Library 2014). The mission of most of these digital repositories is to “enable better access, searchability, usability, and visibility of their research output by those with Internet access” (Ocholla 2011). Although they share a similar mission, there are many avenues to this goal, and repositories vary greatly in their size, content, scope, and successful realization of that mission. As there are many ways in which an IR can engage with the research community, this paper focuses on five areas: interoperability and visibility, engagement and dissemination, researcher participation, education, and information stability.

### **Interoperability and Visibility**

The Association of Research Libraries (ARL) defines repositories as: “institutionally defined, scholarly, cumulative and perpetual, and open and interoperable” (Crow 2002, 3). Most repositories are consistent with those standards if they follow basic OAI-PMH<sup>1</sup> standards for interoperability of metadata. To be truly interoperable, however, repositories need more than comparable metadata. To truly work together, repositories need to be linked to each other in some more meaningful way. For example, IRs should apply more specific metadata to each item so that they are more reliably linked to search engines and can be more accurately indexed and more easily found through the basic searches that are the beginning of so many academic research endeavors (Arlitsch and O’Brien 2012).

Libraries should work to make their sites visible so that the scholarship held in IRs can be indexed by search engines and read and cited in the academic community.

---

<sup>1</sup> [www.openarchives.org/pmh](http://www.openarchives.org/pmh)

That requires time and a specific set of skills. “The goals motivating an institution to create and maintain a digital repository--whether pan-institutional, as a component in the changing structure of scholarly communication, or institution-centric--require that users beyond the institution's community gain access to the content” (Crow 2002). It is not enough to post objects to the internet; the metadata of those objects must be optimized for search engines if these objects are to be found by users.

### **Engagement and dissemination**

Increasing numbers of academic institutions have realized the clear and immediate benefits of having an IR. The “*SPARC Institutional Repository Checklist & Resource Guide*” states that, “institutional repositories offer a strategic response to systemic problems in the existing scholarly journal system—and the response can be applied immediately, reaping both short-term and ongoing benefits for universities and their faculty and advancing the positive transformation of scholarly communication over the long term” (Crow 2002). This is an alluring and impressive draw for libraries to offer scholarly lifecycle literacy education. Scholars see a clear desire for maximum readership, access to and preservation of articles (Chan 2004). An IR could be a good starting place to transform the library into an active participant in the promotion and dissemination of scholarly materials.

IR advocates believe that repositories will make huge changes in the public image of a university’s research output and a library’s image. Crow explains, “the rationale for universities and colleges implementing institutional repositories rests on two interrelated propositions: one that supports a broad, pan-institutional effort and another that offers direct and immediate benefits to each institution that implements a repository” (Crow 2002).

### **Researcher Participation**

It is difficult get researchers to voluntarily submit their own research. It takes time

to explain what the repository is and how it might help a researcher. Researchers need compelling reasons to take time from their busy schedules to assign basic metadata, figure out the copyright status of an article and then find the PDF of the appropriate version of their work. Education and outreach, ease of use, and implied benefits are often not enough to foster a culture of participation that encourages researchers to deposit. Harvard University has implemented a mandate to insure that their repository is populated with its own scholarly output. “The mandate, which resembles a publishing contract, has been instituted to combat rising serials costs that are forcing subscription cuts and restricting intellectual exchange” (Albanese 2008, 16). Even with a mandate, it still takes time and energy to deposit works into a repository.

## **Education**

IRs provide practical opportunities to increase awareness about multiple scholarly communication issues and systems. They also promote “awareness of the increasing importance of virtual, informal, and global mechanisms of communication” (Nabe 2010, 10). They can be a practical platform for education on publishing, copyright, plagiarism, digital preservation and information literacy.

Elsevier's early 2014 takedown notices did more than show the power of the legal department of a major publishing house; they made public the ignorance or disinterest that many published authors have about their copyright holdings (Culter 2013). Researchers share PDFs in an informal network of phone calls, emails and social media (Moriano, Ferrara, Flammini, and Menczer, 2014). As mass dissemination of electronic resources becomes easier, the practice of illegally sharing documents to which authors have signed away the copyright, including distribution rights, becomes easier. Now, instead of illegally sharing an article with a small number of colleagues, an author can share with anyone in the world. This brings two issues to light: one, publishers often have a legal right to these articles, as most publishers still hold the

copyright to the final version of an article; and two, authors do not behave as if they care about that legal distinction. Authors often do not adhere to the copyrights that they agreed to, as seen in the proliferation of illegal copies on services like ResearchGate<sup>2</sup>. IRs can help educate authors about those rights so that they do not sign them away without considering the implications, and when they do sign restrictive author's contracts, they understand what has transpired. Researchers and students should be educated about their rights as authors and the effects that publication agreements have on publications so that they can make more informed choices about their work.

### **Information Stability**

“Link rot” is a term used to describe the permanent unavailability of linked content on web pages due to out of date hyperlinks. The loss of citation information is a real and serious impediment to research. As links deteriorate they no longer point to sources for quotes, maps, data sets and any other digital object that may be deeply important to an article. For example, before publishing a book on the internet, the authors of *Interpersonal Divide: The Search for Community in a Technological Age* attempted to recheck their sources, only to find, years after starting, that 13% of their digital references were gone. “This study continues to challenge advocates of online scholarship to stop touting the convenience of easy access and start resolving issues of later retrieval” (Bugeja and Dimitrova 2011). Online scholarship, they assert, is initially about access to information, but continued retrieval is just as important. Institutional repositories can serve as a permanent source of stable information through intentional preservation architecture and stable URLs. IRs begin to answer the problems in current digital scholarship, although they are not the whole story for most institutions. It is now a question of how to meet this need most efficiently and for the longest amount of time.

---

<sup>2</sup> See: <http://www.researchgate.net>

## **Survey Analysis**

In an attempt to understand the current ecosystem of IRs in the United States the author emailed a questionnaire (see appendix) to 218 managers of the IRs listed on OpenDOAR's list of repositories in the United States. The survey pool was not limited by type or material collected. Twenty-six IR managers responded to the survey, giving a 12% response rate.

Respondents reported the use of several platforms, including DSpace, Digital Commons (BePress), CONTENTdm, a variety of homegrown software, Hydra, Fedora, SobekCM, and debian. Just under half of the responding IRs run on DSpace, a third have contracts with BePress to run Digital Commons, and the remainder are spread among the other platforms. Several types of institutions responded to the survey, including research universities, state archives, a subject repository, teaching-focused colleges, museum libraries, large state schools, and a research institute.

### **Staffing**

Not every library that maintains a repository has an explicitly titled IR manager. Many of the surveyed libraries had a librarian who manages the IR in addition to normal duties. In the most extreme case, a reference librarian was relieved of one reference desk shift in order to run the repository. Some libraries have a team of librarians who work together to run the repository and many have an additional staff person or graduate student. The designation of a single staff person for the IR is a large commitment of resources that not all libraries are able to make to their IRs, but this is not essential for success. Some librarians were concerned that there was someone who ran the IR, yet that person's title did not reflect that role. Some respondents claim 'no real defined leadership' as it is shared among multiple departments. It would benefit the decision-

making process to have a formal plan in place for the staffing and budget of an IR if it is to survive budget cuts, new initiatives, personnel turnover, and the annual assessments of academic institutions.

Most of the institutions surveyed allocated at least .5 FTE to the repository, mostly to education and item deposit. For those libraries that run self-supported repository software (mostly DSpace), technical server and software development and maintenance are not counted in the .5FTE. When making a case for an IR, it would also be prudent to note all the time on the project, regardless of departmental affiliation. This would help administrators and managers see the full extent of work done on and for an IR. Additionally, treating all personnel involved with the IR as valuable players may increase effective communication and simplify management of the project as a whole.

### **Scope of collection**

Most repositories collect Electronic Theses and Dissertations (ETDs), peer-reviewed articles (pre- or post-publication) published in traditional journals, monographs, or book chapters. Additionally, some collect grey literature, technical reports, working papers, white papers, conference presentations, archival university papers, college or departmental newsletters, course catalogs, audio samples, books, journals, accreditation reports, funded grant proposals, artistry and performance materials, video of recorded lectures and talks, university ephemera, conference proceedings, posters, undergraduate work, videos, curricular materials (teaching tools designed by faculty), datasets, images, and maps. One IR currently will add any type of content that supports the university's curriculum and research needs: they collect "everything and anything currently." Some respondents note that there are negative aspects to such a broad collection scope. One of these respondents states: "It's becoming clear that a repository with such a broad scope may be of limited value. Enforcement of metadata best practices is extremely difficult when responsibility for description is distributed to anyone

who requests a collection. Not to mention that lowest-common-denominator description and structure are limiting for projects who want to do more." When asked about aspects of the IR that have been a struggle, this respondent continues:

"The biggest struggle has been defining and maintaining the scope of the repository. This scope has shifted multiple times since starting and new collections/partnerships are given vague expectations about our services. The collection policy is so broad and the scope so varied that providing specialized services per collection is impossible. Another struggle has been running the project without a dedicated budget or business plan. Currently, there are no plans in place to recoup costs of the resources it takes to run the repository (storage, transcoding, streaming, staff, etc.). This makes designing a system with appropriate boundaries and features challenging."

This response clearly illustrates the need for a plan, a collection development policy, and a dedicated, or at least noted, budget. The repository manager knows that these things would help. The struggle is advocating for resources while fighting to keep up with daily operations. The respondent continues:

"Resources that would help: A dedicated budget. A leader with a vision concerning repository services and how they fit into the larger goals of the library/university. Project manager to oversee the development of both systems and digital collections production. More software engineers to work on the platform, new features, and improvements. A repository manager to oversee routine day-to-day queries, data analysis, and technical maintenance Digital archivist to support media collections (standards, policy). More digital lab student staff to grow digitization to support requests. Student staff to support ingestion and description."

One good reason to limit the scope of the collection of an IR to scholarly materials or faculty publications is that it limits the possibilities of mission creep and the

repository, its staff, and budget being pulled in too many directions to be effective at any one task.

### **Age of Institutional Repositories**

The IRs represented by survey respondents had start dates ranging from 1997 to 2013. Although the repositories started at different times, there were similar issues and successes across all IRs, with the younger programs more concerned with faculty engagement and increasing content.

### **Budget**

The survey asked about both the startup budget and estimated annual budget for IRs. The answers varied between hard dollar numbers and guesses. Respondents spent up to \$45,000 for startup with repository services. These repository services average around \$30,000 a year for the survey population. Some respondents reported “there was no dedicated budget” or that “it just came out of the library budget”. It is indisputable that in order to store information, a container is needed in which to house it. These places could be local servers, cloud-based storage, or an external service. In each case there is a cost associated. Additionally, there are personnel costs; these vary from a reassigned 10% FTE to 12 positions that each spend an average of 28% of their time on the IR at one institution. The respondents vary in how they categorized staff who work with the IR—from deposit only to the full range of cataloging, development, server maintenance, liaisons in the institution, marketing, deposits, repository instruction, digitization, and policy development.

“You can't manage what you don't measure,” as paraphrased by Bill Hewlett, has become a management adage that has been applied to information technology, small businesses, corporations, and should also be applied IRs (Symons 2005). The library, an integral part of the research function of a university, should not be driven by bottom line, like a business, but most libraries run on a limited budget, and it would be prudent to

analyze current practices in order to make a case for them if they are ever threatened. Libraries cannot become complacent and assume that because the Graduate School (or similar external funding) gave \$30,000 as startup funds to help maintain the ETDs that they would fund the IR forever. Likewise, when committing to electronic preservation, part of the commitment must be the securing of, or at least planning for, continuing funds. It is difficult to maintain a budget and personnel time without first knowing the amount of time spent currently. "It is self-defeating to advocate for adoption of the service and then be unable to meet the requests of the contributor pool" (Nabe 2010, 12). It is in a repository's best interest to keep good records and maintain a number of metrics to assess the work performed for its upkeep and the benefits enjoyed as a result of that work.

## **Successes**

The survey asked respondents to comment on their biggest accomplishment. Although responses varied from institution to institution, they fell into these categories: ETD work, increasing item counts, publishing journals and student journals, increased outreach, digitization of existing collections, migration to open source software, broadening repository holdings, improved workflows, creating the repository itself, and becoming part of the culture at a home institution. Institutions are increasing the size of their collections and the acceptance and use of those collections. It is difficult to draw conclusions beyond this from the survey responses. Still, it is useful to know what other repositories consider a success.

## **Difficulties**

Difficulties reported in the survey fell into a few categories: lack of time, self-deposit, lack of staffing or a dedicated manager, copyright clearance, obtaining content (especially faculty publications), making time to set policies, IT support, getting better

statistics about use, defining the scope of the IR, funding (especially for storage), managing workflows, tailoring messages to specific faculty, helping other librarians to provide meaningful outreach, and communication and marketing. Though it is difficult to say why these are problematic points for repository managers, the issues can be further broken down into a lack of funding, staffing time, IT support, policies, and outreach. While these are not uncommon for any organization, it is worth noting that these are the places IRs feel pressure.

### **Software issues**

One solution to facilitate the process of building and maintaining an IR was framed as “just purchasing a [commercial product]”. Libraries in this survey who have purchased repository services report similar concerns to their open source counterparts. Both groups spend a great deal of time on outreach, item ingest, and increasing faculty awareness of the tool and service.

Purchasing repository software maintained by a for-profit company could be an easy solution to issues such as technical repository maintenance and gaps in librarians’ skill sets, yet purchasing a product from a for-profit company is at odds with the sustainable continuation of a repository built for the promotion and preservation of open access research.

In an era of shrinking library budgets, committing to preserve the scholarly record without being able to confidently fund the project forever seems tenuous. Many of the surveyed managers reported a current push to have the IR service paid for outside of the library. If the library remains in control of the IR yet not its budget, we take both the infrastructure and the assurance of its continued existence out of our hands.

### **Libraries without Repositories**

Of the libraries surveyed, the few that did not have a repository had all considered an IR at some time. The major reason that they had not pursued an IR was

budget concerns. One library reported that they hoped to join an IR hosted or facilitated by the consortium where they already have membership and working relationships. This is one option that provides fertile ground for collaboration. Another library reported that they had the desire to preserve materials that were produced at the university that could not be supported in their current digital archives. As a small library without a large staff, this library does not have the on-site staff time to create and support an open source product like DSpace for a repository. They also did not have the budget this year to purchase a service like BePress or DSpace Direct. It is this type of issue, caused by limited budget, that collaboration most effectively addresses.

### **Existing Collaboration**

There are a few common existing collaborations within institutions. Just over 37% of the libraries that responded to the survey noted a content or financial relationship with the Graduate School, law school, or individual department. The partnerships include monetary support, outreach and education, and personnel time. Support from departments in the form of items to deposit, mandates, or active makes a large difference in the awareness and population of an IR. These are valuable partnerships that should be fostered on campuses and maintained through mutual benefits.

Many graduate schools support the IR based on the deposit and preservation of electronic thesis and dissertations (ETDs). Two repository managers responded that they had collaborations with other institutions, and many noted that they used services like the DSpace user community to aid their work. These communities and support networks not only distribute work among more people and budgets but they also foster engagement around an IR which is invaluable to its success.

Additionally, sharing the work of an IR within an individual library not only helps the persons responsible for the repository, but as some committees work on things like “tough questions about copyright,” the discussions act as valuable learning tools for all

involved.

## **Potential New Collaboration**

### **Proposal Summary**

The knowledge and resources to run an IR may not be available to all institutions. As librarians we need to think critically about the state of repositories and realize that many situations may not be scalable or sustainable. Although it may be easier for some libraries to divert staff time and resources from the library budget to create a repository, this should not limit the libraries that would like to build a repository yet do not have the resources. For those institutions that may not be able to support that activity but want the benefit of an IR, how can they build a repository that makes sense, is sustainable, and promotes their mission as a library and as an institution?

Five options for collaboration between libraries that would like to establish a new IR are: 1. form a consortium to produce a single repository that has more visibility than any one institution has on its own; 2. increase communication between partner institutions so that we do not all solve the same problems or work in a vacuum; 3. Establish a separate center for site-neutral collaboration, such as OhioLink, OrbisCascade, and OCLC do; 4. share the work of specialists; and 5. use metrics to evaluate and assess IRs so that they may have a productive future.

### **1. Consortium**

For some institutions, even if repositories become less expensive and easier to set up, they are still difficult to maintain and to populate. It may be the case that even in the best circumstances it is advantageous for libraries to join together in this facet of their services. To work with others to produce a single repository that represents many

institutions, and thus has more visibility than any single IR on its own, could be the best option for many potential IRs.

Consortia could be based on any number of attributes of an institution, for example: institutions whose values are aligned, who are similarly sized or focused, who are geographically close, or who are very different and serve to benefit from other's strengths. Existing partnerships or groups (e.g. OhioLINK, the Oberlin group) could strengthen their partnerships and create a shared digital repository. Just as libraries could learn to give up another piece of cataloging control to a body or elected group, they could learn to make decisions about preserving the scholarly record together by working to develop and maintain the repository as a group.

Branding and politics are two issues that institutions may face. Even if the libraries want to work together, it is possible that someone at the Institution will fear the lack of recognition and control of research output. The University of Rochester's UR Research<sup>3</sup> is a good model for the potential success of group repositories. The numerous schools that make up the University of Rochester each have a unique and important identity. Some are famous within their field, e.g., the Eastman School of Music, while some are not; yet, they all share the same homepage. Each body within the university is given space, a logo, and an individual 'search' page within the interface. This allows each school to have its own space and to benefit from the site architecture, metadata, cataloging and maintenance that is specialized and centralized within the library. UR runs the repository on Institutional Repository Software (<http://code.google.com/p/irplus/>), and it serves as a good example of how a consortial repository might break up its content. With more specific design, each school could have its own colors, logo, etc., and make sure that its work was branded. Users could also search by discipline instead of by institutional facet. In a well branded consortial IR the

---

<sup>3</sup><https://urresearch.rochester.edu/home.action>

pages would still be branded so, for example, while on Oberlin College's page a user would still see a Yeoman or the coat of arms at the top of the page and be able to easily navigate to other articles in American Studies from the consortium.

This system could benefit all participating institutions. If an institution is well regarded in the field, it need not worry that joining the open digital research community would detract from it. Instead, the sum of all cooperating institutions would be much greater than any one alone. Even a large research institution would gain site traffic and visibility if allied with another organization. It could be that where individual schools see 'branding', they are really just increasing fragmentation in the digital research world.

Although the costs would be spread over a wide pool of participants, a community repository would still need a more specific funding model. One model would be that each institution pay an even share of costs as yearly membership dues and meet annually to discuss options, changes, and problems as a board. Further, institutions could be contractually obligated to remain in the repository for a minimum of ten years, barring extreme circumstances. This would ensure that no institution had a greater say in running the repository based on funding and that institutions would not be able to withdraw from the collaboration without considerable penalty.

The political aspects of hiring repository managers and the physical site of their workspace would most likely be a tough decision, but one that should be made at the time of hire and set up. The repository could be housed in an existing work space at a participating institution to keep costs down or they could build a new space on neutral ground (see suggestion three). This allows each institution to learn from each other and put pride aside as much as possible for the pursuit of the preservation of digitized knowledge.

## **2. Increased Communication**

At each individual repository, workflows, engagement strategies, budgets, policies, justifications, collection development policies, promotional materials and metadata practices are all duplicated. This does not have to be the case. There are publications and suggestions, yet working in real time with another person to solve problems and make decisions can be extremely helpful. Especially for those IR managers who do not have a committee or working group on site or do not have the time to dig through the literature and forums to find answers to questions, a small body of like-positioned professionals as a support network would greatly reduce the redundancy of effort for IR managers and leave time to do the work of running an IR instead of figuring out how.

### **3. Site-neutrality**

While it may be tempting for institutions to house their resources independently out of a sense of school pride, it is difficult to ignore the benefits of collaboration if it makes these materials more easily found and cited. Increased site activity, improved search index ranking through more robust linking, and a better overall repository experience that results from collaboration can be a good incentive to convince institutions to join together and share the work of a repository. Digital scholarship can be a point of pride, a section of interest on the open web, and uniquely branded, but it does not necessarily need to be uniquely maintained.

Ultimately, a digital repository need not exist in the physical space of an institution. As long as the object pages are appropriately and clearly branded, the information is visible on the open web. With appropriate metadata attached to each item, a server that houses a repository could exist anywhere<sup>4</sup>.

---

<sup>4</sup> There are exceptions to every rule. There are cases where for the protection of sensitive or proprietary data servers are kept locally or in very specific storage spaces. For preservation

One of the benefits of depositing work into a repository is to release knowledge to the global community. Based on this global dispersal, there are few reasons that make local storage important for institutions. One measure of success for a repository is to have documents downloaded as many times as possible. We celebrate our papers making their way to China and Brazil, to Ecuador and France. Items that are ready for that sort of dissemination do not have the sensitivity issues of some data and can easily be stored in servers without concern for geographic location.

An externally based repository could have much the same structure as the consortial system described above, yet this system would be more autonomous. This would set the repository outside of political embattlement and give managers a neutral view over the management of the repository.

#### **4. Shared Expertise**

Many repositories are run on less than 1 FTE, with one person trying to do the work of catalogers, website designers, outreach coordinators, liaisons, coders, copyright experts and still do the rest of their job on top of that. A collaboration allows institutions to share the work of specialists across geographic space so that we do not all have to be experts on each task. If a group of several managers formed together and shared their work and the work of their colleagues, each institution would not have to have a copyright expert, dedicated cataloger, software developer, etc. The work of these specialized employees could be done remotely in exchange for the work of cooperating specialists. Additionally, these specialists could be point people to field questions, design workflows, and troubleshoot problems. Going beyond ‘information sharing’ this model suggests that the actual workloads are shared between institutions.

---

needs it may be prudent to have servers located in disparate places for backup and disaster preparedness.

## 5. Metrics

Repositories that have a clear budget, a dedicated manager, and clear scope of collection development use their resources more efficiently. They are better able to move forward strategically. From an administrative perspective, statistics based on clear metrics help justify budgets (and budget increases when available). In a data driven world repositories cannot afford to not collect data about the services they provide for their institutions.

Thinking seriously about the worth of a repository requires measures and indicators that guide that process year to year. Use is often a key indicator of the success of an institutional repository. Use is measured in downloads, views, and citations. These statistics measure only one aspect of a repository. The following metrics represent some additional indicators of the success of IRs.

**Number of Items:** One of the major metrics that give meaning to a repository is its contents. Many repositories set a scope of collection to define the types and kinds of items that it may ingest. It is prudent to measure not only the quantity and quality of work in the repository, but also the growth over time. Size of a collection should be measured in comparison with the total output of possible items.

**Visibility and Indexing:** Another important metric for IRs is visibility to search engines. If appropriate metadata is not applied, our content will not be visible from the various search engines and open access aggregators and directories that work to make this information more prominent on the web. According to the Jisc Infonet Repositories Toolkit, a good deal of OA IR content can be found through basic internet searches. This is only true, however, if the correct metadata (Open Archives Initiative Protocol for Metadata Harvesting [OAI-PMH]) is applied, the content is held on a reliable server, and the HTML or microdata is optimized so that search engines can more easily index our IR

content.

A number of services have been developed to enable more specific content searching and thus more relevant results than are possible from general online search engines<sup>5</sup>. These specific searches are not, however, what the majority of people use when searching for information. It is important to also optimize metadata for general search engines, like Google Scholar, using additional meta tags and markup, so that these services can index information and provide accurate citation data to the reader. A metric to measure this would be the ratio of indexed items in an IR to the total number of items in the IR.

**Awareness:** If more people know what an IR is and that it exists at an institution, then more conversation about the particulars of copyright, logistics, and item deposit can occur. As awareness grows, there will be increased conversation on campus about scholarly communication issues, which will increase information seeking and understanding of the topic. Awareness on campus can lead to appreciation at an institution, which is valuable for long term administration of a service. This can be measured in polls, in site visits, in reference questions on the topic, or through pulse checks when visiting departmental meetings.

**Understanding:** In order to measure the impact of a repository as an educational tool, it is useful to define a population, a set of skills or knowledge, and a measure of that knowledge. The groups that should be educated include, but are not limited to, faculty (especially those who publish), graduate students, undergraduate students, and administrators. Each group has a specific subset of scholarly communication education that would be the most valuable to their work. Repository

---

<sup>5</sup> Access Options include: BASE (<http://www.base-search.net/>), OAIster (<http://www.oclc.org/oaister/>), OpenDOAR Search (<http://opendoar.org/search.php>), and DART-Europe (<http://www.dart-europe.eu/basic-search.php>)

managers should administer surveys, run brief interviews, and listen to the conversations around campus to gauge the knowledge base at their institutions. Once there is a baseline of knowledge on campus, IR managers can reach out to the community in a strategic way and develop targeted, measurable education activities about authors' rights, contracts, and copyright. The number, type, and impact of each teaching situation, if measured, can become a valuable metric when making a case for the tangible use of a repository beyond storage.

**User Experience:** User experience can also be used as a metric to judge the success of a repository. The user end of the repository should be easily accessible and provide a clear search and retrieval experience. The institution should plan for at least one initial usability study, with the potential to complete additional usability studies for the duration of the project.

## **Conclusion**

Based on this survey of IR managers, there is a wide range of issues and barriers that inhibit repositories from optimal performance or planning for a sustainable future. Though 37.5% of repository managers surveyed reported some existing collaboration, many of the partnerships are within a single campus. These partnerships exist mostly between institutional bodies that have a research and scholarship interest, i.e., the library, the graduate school, the research office, a law school, or academic department. While a few inter-institutional collaborations already exist, there are many more connections that can be developed to benefit repository managers and the ecosystem of scholarly work housed within repositories. New collaborations could be centered on a shared repository, shared workflows and increased communication, shared facilities and infrastructure, shared personnel and expertise, or a shared

understanding of how to best measure the successes of an institutional repository. There is a healthy and growing ecosystem of repositories in the United States and globally upon which to build beneficial collaborations. By working together, we lessen the burden on each institution. Repositories that are built and maintained using a sustainable and collaborative model will benefit more libraries in the present and more users in the future.

## References

- Albanese, A. 2008. "Harvard Mandates Open Access." *Library Journal* 133(3): 16-17.  
<http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=502942865&site=ehost-live>.
- Arlitsch, Kenning, and Patrick S. O'Brien. 2012. "Invisible Institutional Repositories: Addressing the Low Indexing Ratios of IRs in Google Scholar." *Library Hi Tech* 30(1): 60-81.
- Bankier, J. and Smith, C. 2010. "Repository Collection Policies: is a Liberal and Inclusive Policy Helpful or Harmful?". *Australian Academic & Research Libraries* 41(4): 245-259.
- Bugeja, Michael J. and Daniela V. Dimitrova. 2011. *Vanishing Act; the Erosion of Online Footnotes and Implications for Scholarship in the Digital Age*. Sacramento, CA :Litwin Books, LLC.

- Chan, L. 2004. "Supporting and Enhancing Scholarship in the Digital Age: The Role of Open Access Institutional Repository." *Canadian Journal of Communication* 29(3): <http://cjc-online.ca/index.php/journal/article/view/1455/1579>
- Cornell University Library. 2014. "*arXiv Membership Model*.  
<http://arxiv.org/help/support/faq>.
- Crow, Raym. 2002. "The case for institutional repositories: A SPARC position paper." *ARL: A Bimonthly Report on Research Library Issues & Actions* 223: 1-4.  
[http://sparc.arl.org/sites/default/files/media\\_files/instrepo.pdf](http://sparc.arl.org/sites/default/files/media_files/instrepo.pdf).
- Crow, Raym. 2002. "SPARC Institutional Repository Checklist and Resource Guide." The Scholarly Publishing & Academic Resources Coalition.  
[http://sparc.arl.org/sites/default/files/presentation\\_files/ir\\_guide\\_\\_checklist\\_v1.pdf](http://sparc.arl.org/sites/default/files/presentation_files/ir_guide__checklist_v1.pdf)
- Culter, Kim-Mai. 2013. "Elsevier's Research Takedown Notices Fan Out To Startups, Harvard, Individual Academics." *Techcrunch*  
<http://techcrunch.com/2013/12/19/elsevier/>
- Gilman, I. K., Gilman, M. I., & Kunkel, M. 2010. "From Passive to Pervasive: Changing Perceptions of the Library." *Collaborative librarianship* 2(1): 22-32.  
<http://web.ebscohost.com>.
- Google Project Hosting. 2014. "Irplus Institutional Repository Software."  
<http://code.google.com/p/irplus/>.
- Johnson, Richard K., 2002. "Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication." *D-Lib Magazine* 8:11.
- Moriano, Pablo, Ferrara, Emilio, Flammini, Alessandro, and Menczer, Filippo. 2014. "*Dissemination of scholarly literature in social media*." figshare.  
<http://dx.doi.org/10.6084/m9.figshare.1035127>
- Nabe, Jonathan A. 2010. "*Starting, Strengthening, and Managing Institutional Repositories*." New York: Neal-Schuman Publishers, Inc.,.

National Center for Biotechnology Information. 2011. "PMC Overview."

<http://www.ncbi.nlm.nih.gov/pmc/about/intro/>.

Ocholla, D. N. 2011. "An Overview of Issues, Challenges and Opportunities of Scholarly Publishing in Information Studies in Africa." *African Journal Library, Archives & Information Science* 21(1): 1-16.

OCLC. 2014. "A History of Cooperation; A Future that Depends on it."

<http://www.oclc.org/about/history/default.htm>.

Oder, N. 2009. "Study: In Downturn, Academic Libraries to Focus on Value, ROI."

*Library Journal* <http://www.libraryjournal.com/article/CA6705634.html>

OhioLINK. 2014. "Welcome to Ohiolink." <http://www.ohiolink.edu/>.

Open Archives Initiative. 2014. "Open Archives Initiative Protocol for Metadata Harvesting." <http://www.openarchives.org/pmh/>.

Symons, Craig. 2005. "IT Strategy Maps: A Tool for Strategic Alignment Visual Representation Of Strategy Makes IT-Business Communication Easier." Forrester Research <http://cendoc.esan.edu.pe/fulltext/e-documents/ITStrategyMaps.pdf>

Cybermetrics Lab. 2014. "The Ranking Web of World Repositories- North America" <http://repositories.webometrics.info>.

University of Rochester. 2014. "UR Research."

<https://urresearch.rochester.edu/home.action>.

## **Appendix**

### **Questionnaire**

1. What is the name of your repository?
2. What platform do you run your repository on?
3. Who leads the work of your IR and what is their title?
4. Is there a committee or advisory group in place to support the work of your IR?
5. What kind of materials do you collect?(Scholarly, archival, audio, visual materials, etc.)
6. When was your repository started?
7. What was the startup budget?
8. What is your estimated annual budget?
9. What percent of personnel assignments and time are spent on the repository?
10. What do you and your colleagues spend the most of your time on?
11. What is your biggest accomplishment? What helped that to happen?
12. What has been a struggle? What resources could help solve that?
13. Have you had any support from other departments, universities or other outside organizations? If so, how were you supported?

14. Is there anything else you would like to share about your repository or the work you do in support of it?