# Topic 1
# All Universities Should Have an Institutional Repository

## Institutional Repositories: The Great Debate

### Affirmative Argument

**Soo Young Rieh**
Associate Professor, School of Information, University of Michigan
Email: rieh<at>umich.edu

The MIRACLE (Making Institutional Repositories a Collaborative Learning Environment) Project team at the University of Michigan reports that about 50% of four-year colleges and universities in the United States have either planned or implemented institutional repositories [1] while the other half has done no planning for institutional repositories. University administrators, library administrators and staff in these non-planning institutions may, however, need to decide whether to jump on the institutional repository bandwagon. This article's purpose is to convince decision-makers in the non-planning stage that it is time for them to act on developing an institutional repository.

### "Why don't you just do this?"

One librarian who participated in our case studies said that she got a clear message from an institutional repository (IR) manager in one of her peer institutions: "Why don't you just do this?" Even though "this" goes by various different names – *DSpace* (MIT), *Deep Blue* (Michigan), the *Knowledge Bank* (Ohio State), *ScholarWorks* (University of Massachusetts), *IDEALS* (Illinois) and *eScholarship* (California), these repositories have similar purposes and functionalities across institutions. Library staffs, library directors, archivists and others involved in institutional repositories possess a good understanding of the kinds of benefits and value an institutional repository can bring to both academic institutions and society [2].

First, an institutional repository provides the opportunity to create one

### Negative Argument

**Kevin Smith**
Scholarly Communications Officer, Duke University
Email: Kevin.l.smith<at>duke.edu

When I was a child, one of the favorite places for my brother and me to play was in the attic of my grandmother's house. There were a couple of specific attractions of her attic for two young and mischievous boys. First, the attic seemed to be a hidden place, since the doorway to its staircase was in a closet and behind a rack of clothes in large storage bags. Second, we never knew just what we would find up there. I remember vividly an old-fashioned electric fan that worked only sporadically, several incomplete jigsaw puzzles, pieces of antique luggage covered in destination stickers and a GI Joe figure that was missing one boot. Lots of fodder for imagination and play, but, since things in the attic were pretty well abandoned, little effort was made to keep them complete and in working order.

My grandmother's attic was a wonderful place to play, but it is a poor model for an institutional repository. I have a very real concern that, if the proposition were adopted, many of the resultant repositories would look a lot like that attic.

First, as repositories proliferate, especially at smaller institutions that may lack adequate funding, staffing and expertise, the issue of finding material that is hidden in those repositories becomes a concern. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) [5] provides a structure for creating repositories that can be searched by Google and made interoperable with other repositories and search tools,

RIEH, continued

Affirmative, from page 12

central virtual place into which university members can deposit their scholarly and administrative digital content. When various digital materials, including faculty e-prints, student work and archival primary sources, are put together, digital content has the potential to become the greatest intellectual capital of an institution. Once a digital collection is in place, excitement quickly mounts as it introduces a whole new way of accessing and using digital content. As one librarian said during an interview, an IR "provides access to those collections that no one would ever know exists." Another library staff member sounded even more excited in claiming that an institutional repository allows "serendipitous discovery across disciplines that was not possible."

If all this enthusiasm does not sound convincing enough, there is more. Given that technology will only keep changing, individuals, though able to keep up with the technology in their working and everyday lives, may well find themselves unable to migrate their own personal digital content from one technology to another without a system – like an institutional repository – that makes it possible to manage technological changes institutionally while promising to preserve intellectual output in the long term.

### "We are always interested in what our peer institutions are doing."

*The Census of Institutional Repositories in the United States* [1] found that there is "a sleeping beast of demand" on the part of small and mid-size universities and colleges with respect to institutional repositories. Although the majority of small and mid-size institutions have not yet begun planning for an institutional repository, they proved surprisingly positive about institutional repositories and were certainly interested in "stories about how small institutions made their institutional repositories a reality." These institutions have been somewhat slow in joining the institutional repository bandwagon not because they were unaware of the value of IRs but because they lacked models, best practices and guidelines from their peer institutions. In fact, their institutional repositories look very different from those at research universities regarding content, audience, qualifiers and even uses. For instance, small and

SMITH, continued

Negative, from page 12

but its application requires a degree of sophistication that may not be available at every university.

At this point it is worth a digression to consider the definition of a university in the proposition that all universities should have an institutional repository. If *university* is defined quite narrowly to include only very large institutions, some of the problems about staffing and funding may not be too serious; however, such a constricted definition is problematic and artificial. There are many smaller universities and colleges (the distinction is often blurry or simply not observed by schools that retain traditional labels as curriculums expand or contract) that have unique materials and faculty scholarship that deserves exposure in a digital repository. Subjecting those institutions to an arbitrary requirement that each have their own such archive could either discourage them altogether or lead to inexpertly managed repositories. One of the potential problems would be a proliferation of non-OAI compliant repositories with content that would be very difficult to locate. Although some schools might want a repository for purposes of "branding" themselves, we need to remember that very few searchers access our digital content by "walking through the front door." If repository content cannot be found through traditional means of searching the Internet, that content will be as hidden as the stairway to Nana's attic was.

The second characteristic of that attic was that it contained a motley and unsorted collection of stuff. If an institutional repository is OAI compliant, and users find material by searching in Google Scholar or a similar tool, the random nature of deposits in small repositories will not be visible or problematic. Conversely, if users of a non-OAI compliant repository were to come through the branded portal, the thin and unselected nature of the materials they might encounter would be obvious.

There are several ways to address this potential embarrassment for institutions lacking the necessary resources to build and maintain strong IRs. One is to take advantage of the editing and peer-review processes

RIEH, continued

**Affirmative,** from page 13

mid-size institutions are often more interested than large institutions in consortia. Also, the institutional repositories of these smaller institutions may have digital content oriented more toward teaching objects than research products.

These small- and mid-size institutions recognize the value of students' work as an important part of institutional repository content. A library director in a small university reported during an interview that her institution saw an institutional repository as "something that's potentially positive because it's a way for students to get their work out and around beyond the boundaries of themselves and their professors." An archivist at a small college expressed excitement on realizing that an institutional repository provides a great opportunity to archive student newspapers.

Interview participants in the MIRACLE Project apparently agree that preservation is "one of the biggest things" across various types of universities and colleges. While there are still many unknowns in building and maintaining digital collections in perpetuity, most people involved in institutional repository development are confident in the long-term sustainability of institutional repositories. The good news is that the confidence level is even higher at institutions that already have operational IRs than at those in which institutional repositories are still being planned or pilot-tested [3].

### "I don't know where I could archive it."

The discussion above focuses on the perceptions of library staff and administrators based on interview data collected during the MIRACLE Project's phone interviews conducted in the fall of 2006. Next, we consider the perspectives of institutional repository users, of which there are two types: contributors and searchers.

An example of a contributor is a scholar who has a dataset from his research projects, but doesn't "know if I should be saving [this], I don't know how I would describe [it], I don't know where I could archive it, I don't know if I should be sharing it or how I could share it – could you help me figure this out?" [3] Institutional repositories are not the only type of self-archiving venue: 684 respondents of Kim's (2008) survey reported self-archiving in personal

SMITH, continued

**Negative,** from page 13

that most scholarly journals already have in place by limiting IR deposits to articles accepted for publication in such journals. Such a policy would preserve a level of quality, but it would exclude potentially valuable material like archival materials, working papers, data sets and the like, as well as most multimedia objects. Another option would be to have an editorial process in place, by which submissions to a repository would be evaluated, corrected and sorted into appropriate collections of similar material. To do this task, of course, the repository would need a dedicated and professional staff and a critical mass of content – both of which are unlikely for many IRs at smaller universities. Finally, a small university could elect to combine its content with that of other institutions, either on a disciplinary basis or within a consortium, in order to reach critical mass and obtain some level of professional management. OAI compliance would be necessary for this cooperation to happen, but the technical and metadata expertise necessary could be pooled across the consortium. This option, I argue, is the real solution for many universities and is the reason the proposition that all universities should have their own repositories must be rejected.

A final similarity between some repositories and my grandmother's attic is that the material contained in each may be left to decay, deteriorate and be forgotten. For an institutional repository to serve its important functions, there must be a commitment to preservation that is simply beyond the means of many small institutions. At this point in time, we just do not know how long digital objects will persevere without intervention. It is also impossible to predict accurately the changes in technology that will make one format inaccessible and necessitate a transition to some new, yet-to-be-created file structure. A real promise of preservation requires support for an open-ended commitment to necessary future cost, which remains unknown and, probably, unknowable. Even large universities doubt whether this commitment is sustainable. It would be foolish and irresponsible for smaller or less well-funded institutions to jump into the institutional repository business without any hope of living

RIEH, continued

**Affirmative,** from page 14

web pages, department/school/college websites, research group/lab/center websites, disciplinary repositories and/or institutional repositories. Although institutional repositories were not the top choice for most self-archiving researchers, they can be appealing, given their ability to preserve materials [4].

Now let's turn to the other type of institutional repository users – searchers. Why do they want to use the institutional repository? The institutional repository searchers whom the MIRACLE Project team interviewed stated a wide variety of benefits associated with using institutional repositories. For instance, one interviewee asserted, "It's kind of like a central receptacle of research going on only here and it's worthwhile to get an idea what's around you before you search on a search engine that's much broader." Another way in which institutional repositories are outshining other information systems is their perceived trustworthiness. Many of our interviewees expressed the opinion that institutional repositories and their content are more reliable than information from search engines such as Google and Google Scholar.

## All universities should have an institutional repository.

Nobody can say for certain how an institutional repository should be developed, what kinds of content should be included or how the system should be maintained. However, this much is certain: if an institution has not undertaken planning an institutional repository, it should consider doing so now. Even though there is quite a bit of uncertainty remaining about how institutional repositories should function, joining the institutional repository bandwagon right now is important if an institution does not want to fall behind. A library director who participated in our case studies said, "Capacity is the best way to think about it [institutional repository]." That point is well taken. Institutional repositories should be considered an infrastructure investment for the future. When self-archiving becomes an established norm in the scholarly and social culture, universities should already be prepared with a technologically supportive mechanism. All universities should be proactive in developing a strategy for scholarly publishing and preservation. Institutional repositories will, without doubt, play a key role in evolving in that direction. ∎

SMITH, continued

**Negative,** from page 14

up to the necessary commitment to sustainability and preservation. Benign neglect was sufficient for our attic-cum-playground, but it is not a workable model for a functioning institutional repository.

For institutional repositories to become a sustainable presence and an agent of change on the landscape of scholarly communications, they must grow out of carefully considered decisions universities make on the basis of local needs, a sensible commitment of resources and a rational awareness of the long-term commitment to support. For some institutions, that decision will inevitably be in the negative. For those universities it is important to provide alternatives that can help them expose valuable assets that contribute to scholarship without burdening them with guilt or negative judgments. Shared repositories may provide a good solution for these institutions. Such repositories may be disciplinary, modeled on the very successful archives already in operation for physics, economics and other disciplines. (For example, see the ArXive repository for physics, mathematics, computer science, quantitative biology, quantitative finance and statistics at http://arxiv.org/, and the RePec collaboration in economics at http://repec.org/.) Alternatively, they may be consortial efforts such as the ones managed by the OhioLINK consortium, for example, their shared repository of electronic theses and dissertations at www.ohiolink.edu/etd/

These shared solutions to open access for scholarly content offer the best opportunity for repositories that are high quality, sustainable, and in which important material can be located easily. Such cooperation is not merely for the small institutions that cannot afford their own single-institution repositories; larger universities may find that cooperative development leads to better repositories that offer both a critical mass of content and a stronger trajectory into the future. Insofar as the proposition with which this debate begins – that all universities should have an institutional repository – downplays the importance or hinders the development of such shared solutions, it must be rejected. ∎

Resources Mentioned in the Topic 1 Debate

[1] Markey, K., Rieh, S. Y., St. Jean, B., Kim, J., & Yakel, E. (2007). *Census of institutional repositories in the United States: MIRACLE Project research findings.* (CLIR Publication No.140). Washington, D.C.: Council on Library and Information Resources. Retrieved February 23, 2009, from www.clir.org/pubs/reports/pub140/contents.html.

[2] Rieh, S. Y., Markey, K., Yakel, E., St. Jean, B., & Kim, J. (2007). Perceived values and benefits of institutional repositories: A perspective of digital curation. *An International Symposium on Digital Curation (DigCCurr 2007), Chapel Hill, NC, April 18-20, 2007*. Retrieved February 23, 2009, from www.ils.unc.edu/digccurr2007/papers/rieh_paper_6-2.pdf.

[3 Rieh, S. Y., St. Jean, B., Yakel, E., Markey, K., & Kim, J. (in press). Perceptions and experiences of staff in the planning and implementation of institutional repositories. *Library Trends*.

[4] Kim, J. (2008). *Faculty self-archiving behavior: Factors affecting the decision to self-archive.* [Unpublished] doctoral dissertation, University of Michigan.

[5] Open Archives Initiative. (2002, June 14). *Open Archives Initiative protocol for metadata harvesting, v. 2.0 (OAI-PMH)*. (Document version 2008-12-07T20:42:00Z). Retrieved February 23, 2009, from www.openarchives.org/OAI/openarchivesprotocol.html

16