February 2009

I Hear the Train A Comin' -- International Repositories: The Promise of Yesterday

Greg Tananbaum

Scholar Next, gtananbaum@gmail.com

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Tananbaum, Greg (2009) "I Hear the Train A Comin' -- International Repositories: The Promise of Yesterday," *Against the Grain*: Vol. 21: Iss. 1, Article 53.

DOI: <https://doi.org/10.7771/2380-176X.2536>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

I Hear the Train A Comin' — Institutional Repositories: The Promises of Yesterday

Column Editor: **Greg Tananbaum** (Consulting Services at the Intersection of Technology, Content, and Academia) <gtananbaum@gmail.com> www.scholarnext.com

I recently had the privilege to address the **Association for Library Collections and Technical Services (ALCTS)** at their Midwinter Symposium. The topic at hand was institutional repositories, specifically their benefits and challenges. The **ALCTS** team asked if I could frame the session by talking first about what we have learned to date from the IR experience, then speculate on where we are headed. This column, adapted from my talk in Denver, focuses on the reflective piece of that equation. It focuses on the initial promises of the IR concept, and how well reality has matched these expectations.

The first glimmer of what would become today's institutional repository flickered to life in 2000 with the launch of the **ePrints** platform, built at the **University of Southampton** off of the **CogPrints** chassis. I think most folks will agree, though, that the IR Era began in earnest in 2002 with the launch of the **DSpace** and **Digital Commons** platforms. **DSpace** and **Digital Commons** made institutional repositories appealing, albeit in different ways. **DSpace** rendered a do-it-yourself IR solution a truly attractive possibility for the first time. This was due in part to the strength and flexibility of the software, and in part to the validation that **MIT's** and **HP's** interest gave the notion of institutional repositories. **Digital Commons** demonstrated that the private sector thought there was something to the IR concept. It also threw the weight of the **University of California** — **bepress's** first customer — behind the push to promote IRs as an important development in the scholarly communication realm.

Soon enough, lots of institutions were experimenting with **ePrints**, **DSpace**, **Digital Commons**, and the handful of other systems that popped up. By 2005, the **Directory of Open Access Repositories** included 500 IRs worldwide. Today that number exceeds 1,300. Why do so many institutions have repositories? What do they want from them, and are these expectations being met?

From 2002 onward, institutional repositories have tantalized the scholarly communication space with the following six promises:

1. IRs are a concrete response to scholarly communication crisis.

I won't spend any time reviewing the scholarly communication crisis of the 1990's onward. Suffice to say, libraries were asked to pay a lot more for access to information. Institutional repositories were a tangible way that the library could, at least on some level, strike back. Schools could use their repositories to disseminate intellectual outputs directly to the world without the publisher as intermediary. In some ways, this was a warning shot across

the bows of publishers that if the perceived predatory pricing practices did not change then universities might increasingly take matters into their own hands.

The IR also showed potential in staunching the long decline of certain types of scholarly materials that were also in various forms of crisis. Here I am thinking most obviously about monographs, which from the late 1980's forward have hovered on or near the endangered list. But I also have in mind niche materials like handbooks, *Festschriften* and white papers. The IR provided a glimmer of hope in the form of a low-cost mechanism to produce and disseminate these content forms.

2. IRs expand access to scholarly information.

Beyond traditional peer-reviewed publications, the Internet showed us that less formal, more in-the-moment communications might have value. The institutional repository became seen as a great place to collect this grey literature under one roof, as it were. In doing so, the hope was that the materials would be more easily discoverable to a wide audience — there would be less need to search through the various departmental and personal Web pages where these materials were previously likely to be stored.

Related to this notion of expanding access to a broader array of scholarly communication forms was the hope that institutional repositories could provide a home for data sets, multimedia files, and executables. The emergence of the IR category coincided nicely with the rise in bandwidth capabilities and the drop in storage costs, making it relatively easy to transfer and manipulate big, fat files. The IR was viewed as a mechanism to categorically collect, organize, and disseminate this type of information.

Regardless of the content type, an increasing concern as scholarly materials migrated to the Web was that their availability would be evanescent. The 404 Not Found error was particularly galling to librarians who valued long-term information accessibility. By housing this information within the institutional repository, the hope was that its long-term availability would be assured.

3. IRs highlight the depth and breadth of the institution's intellectual output.

Another promise of the institutional repository was that it would be a good promotional tool for recruiting students and faculty, for fundraising, for alumni relations, and for profile-building among the general public. The IR could show the world what a school was up to, what its organized research units were all about, how innovative individual faculty members were, and how the school was

taking the lead in specific subject areas. Rather than a hodge-podge of individual departmental and personal pages operated by graduate students and potentially disgruntled IT personnel, the IR could paint these pictures more efficiently by organizing information, maximizing accessibility, and building a UI that supported the marketing mission.

4. IRs have an accelerating effect on the "information wants to be free" movement.

The evolution of the institutional repository corresponds roughly with the rise of open access, and there are some obvious philosophical overlaps. Both IRs and OA seek to eliminate the access costs — namely, subscription fees — associated with traditional scholarly communication. Both seek to leverage technology, specifically the lowered operational and dissemination costs associated with electronic workflows, to reduce the production costs associated with traditional scholarly communication. Some within the scholarly communication space felt that IRs were a nice complement in the sense that they were another visible, understandable challenge to the traditional model of publishing. The thought was that the IRs imprimatur might perhaps take the place of the traditional publishers' seal of approval.

5. IRs are a potential breeding ground for a new generation of university-founded e-journals.

Very much related to #4 was the notion that the institutional repository contained much of the editorial and production mechanisms necessary to operate e-journals. The IR could, as a result, make it much easier to launch new journals without aid of a commercial or society publisher. There is a corollary here that the IR and the various publishing opportunities it provided might prove useful in helping the institution retain some of the intellectual capital produced under its auspices rather than ceding it to other publishing outlets as was the norm.

6. IRs require low adoption costs for authors.

The sixth promise of the institutional repository was that it would be straightforward to secure faculty participation. The software was simple to use. Posting took but a few minutes. The small hassle of posting was more than outweighed by the wider dissemination of content, its long-term accessibility, the so-called "rising tide benefit" of being part of an expansive and important collection, and the elimination of the need for clunky personal or departmental pages. The ease of author adoption would mean that libraries could reasonably and effectively run point on the marketing aspects of IRs.

continued on page 91

I Hear the Train A Comin' from page 90

What follows is a wholly subjective assessment of how fully the institutional repository has lived up to these six initial promises. As a concrete response to the scholarly communication crisis (Promise #1), IRs have had some tangible influence. They have put the university in a proactive stance as it explores ways to lower the costs associated with information dissemination. The IR has pushed publishers to change their policies in relation to the free dissemination of at least some version of a published article (typically a postprint). The institutional repository has also proven to be a viable home for the monograph, the edited volume, and other niche publications. While the IR has not conclusively ended the scholarly communication crisis, it certainly has helped to mitigate some of the more troublesome aspects of it.

In terms of expanded access to scholarly information (Promise #2), institutional repositories have had mixed success. On the plus side, IRs have delivered eyeballs to the literature. **Bepress's ResearchNow** database, a sort of meta-repository site that allows searching across the 70 or so **Digital Commons** installations, has logged 10 million full-text downloads in the last 365 days. The vast majority of these are for working papers and other non-journal content. IRs have simply made this type of content more easily discoverable. Also on the plus side, these materials are substantially more likely to be locatable and accessible in five or ten years' time when compared to the old tangle of personal and departmental Web pages. Librarians care a lot more about curating and archiving than any other stakeholder, and their oversight of the repository provides a healthy guarantee that IR deposits will not be 404'd someday soon.

Now the negative. One of the big hopes of the IR — that it could be a mechanism to categorically collect, organize, and disseminate data sets, multimedia files, executables and other non-static information — has not been realized in any substantive way. It is undoubtedly true that we have seen some interesting experiments along these lines, but the IR has not rendered the dissemination of non-static materials a commonplace occurrence. The institutional repository has made this technically possible, but has thus far failed to make much progress beyond that.

Institutional repositories have done a fair job of highlighting the depth and breadth of the institution's intellectual output (Promise #3). Many schools have placed special emphasis within their repositories on specific subjects or programs where they excel. The IR has proven to be a good central depository for collecting disparate materials that emphasize a school's leadership in a certain subject. Examples include **Boston College's Church in the 21st Century Series** and **Cornell's Industrial & Labor Relations Collection**. On the down side, I am not aware of many, if any, instances where an institution has coordinated its IR activities with an alumni association or a fundraising drive or a central campus PR

campaign. The institutional repository may be a treasure trove of valuable information demonstrating leadership and innovation in specific discipline, but it is unclear to me if anything is being done to systematically leverage this golden asset. The IR has potential as a marketing tool, but I don't believe that potential has been anywhere close to fully realized as yet.

One of the undeniable successes of the institutional repository has been in furthering the "information wants to be free" agenda (Promise #4). There is simply a lot more content that is more readily available than there was in 2002. The **OAIster** database contains nearly 20 million records at this point. Not all of them are IR materials, but a lot of them are. Institutional repositories have helped create an expectation among researchers, particularly younger ones, that some form of the materials they seek may be freely accessible to them with a modest amount of Web exploration. I suspect this will be one of the more lasting impacts of IR on the scholarly communication realm.

The IR has in many cases lowered the barriers to the launch of new e-journals (Promise #5). **Digital Commons**, for example, has been used to produce close to 150 open access e-journals. That is an impressive number, roughly on par with the number of titles **Hindawi** publishes. However, the reduction in cost and effort it takes to start a journal is not all to the good. It encourages vanity publications, half-hearted endeavors, and other projects that are likely to add clutter rather than clarity to the scholarly communication picture. In addition, many IR-driven journal launches

seem to have taken place in a vacuum. There has been little coordination with other campus units, including but not limited to the university press. The result is that we see random buds sprouting across the scholarly terrain as opposed to a well-tended and planned garden.

The institutional repository has been a disappointment in terms of the adoption costs for authors (Promise #6). The software has indeed proven relatively simple to use. Posting does only take a few minutes, as the advocates had promised. However, scholars have by and large been unconvinced that the effort of posting is outweighed by the benefits of wider dissemination, long-term accessibility, and so forth. Content acquisition has been a slog. This has forced the library to be more creative and aggressive in its marketing efforts, a task for which the library is not ideally suited.

So this is how I see the IR world reflected in the rear view mirror. We have done some things well and missed the mark in other areas. In next issue's column, I will address the key benefits institutions can recognize via a successful repository, as well as the possible impediments to a successful IR that institutions must face. I'll also look at the future of institutional repositories within the larger context of a rapidly changing scholarly communication landscape. Stay tuned. 🌸

Technology Left Behind from page 89

by some of the resources highlighted above, it is not always necessary to spend a lot of money to meet the user's need. The important part is to meet the need.

Web Resources

Aquabrowser — <http://www.aquabrowser.com/>
Bibster — <http://bibster.semanticweb.org/>
BiblioCommons — <http://www.bibliocommons.com/>
Cherokee County Public Library on LibraryThing — <http://www.librarything.com/profile/cherokeelibrary>
Delicious — <http://delicious.com/>
Discogs — <http://www.discogs.com/>
Flixster — <http://www.flixster.com/>
LibraryThing — <http://www.librarything.com/>
Oakville Public Library — <http://www.opl.on.ca/>
PennTags — <http://tags.library.upenn.edu/>
WorldCat — <http://www.worldcat.org/>

References and Further Reading

Farkas, Meredith. "Isn't it Delicious?" *American Libraries* 39, no. 4 (April 2008): 32.

Fonticharo, Kristin. "Is LibraryThing Actually

Useful in a Library Media Center?" *School Library Media Activities Monthly* 24, no. 5 (January 2008): 28-29.

Furner, Jonathan. "User Tagging of Library Resources: Toward a Framework of System Evaluation." *ICBC* 73, no. 3 (July/September 2008): 47-51.

Hadro, Josh. "Tagging Added to WorldCat.org." *Library Journal* 133, no. 17 (October 15, 2008): 21. <http://www.libraryjournal.com/article/CA6602838.html?q=tagging+worldcat>.

Oder, Norman. "BiblioCommons Emerges: 'Revolutionary' Social Discovery System for Libraries." *Library Journal*. <http://www.libraryjournal.com/article/CA6579748.html>.

Oder, Norman. "BiblioCommons: A Breakthrough?" *Library Journal* 133, no. 2 (August 15, 2008): 14. <http://www.libraryjournal.com/article/CA6585865.html?q=bibliocommons>.

Robbins, Elaine. "Enhancing the OPAC through Library Thing." *Against the Grain* v.19#6 (December 2007/January 2008): 72.

Sanders, Dinah. "Tag - You're It!" *American Libraries* 39, no. 11s (December 2008): 52-54. 🌸