

6-6-2008

Institutional Repositories—Strategies for the Present and Future

Connie Foster

Western Kentucky University, connie.foster@wku.edu

Follow this and additional works at: http://digitalcommons.wku.edu/dlts_fac_pub



Part of the [Science and Technology Studies Commons](#), and the [Technology and Innovation Commons](#)

Recommended Repository Citation

Foster, Connie, "Institutional Repositories—Strategies for the Present and Future" (2008). *DLTS Faculty Publications*. Paper 4.
http://digitalcommons.wku.edu/dlts_fac_pub/4

This Article is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in DLTS Faculty Publications by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.

Institutional Repositories—Strategies for the Present and Future

By Connie Foster

Professor and Head, Department of Library Technical Services

Western Kentucky University

Bowling Green KY 42101

NASIG Conference, June 6, 2008

Have you ever searched a departmental web site of a prolific science or psychology faculty member who receives frequent accolades for presentations and publications? Can you find their research and link to it, or even pinpoint their recent articles under departmental listings? Results of this endeavor can be quite surprising. Often departmental or faculty web pages reveal what is *not* there, how hard it is to find current citations and works of a professor, much less their entire body of work. While an institutional repository (IR) will not solve all of these access and retrieval difficulties, it can offer a way to bring together much of the intellectual and creative efforts of a university in one place and establish a permanent path to discovery and open accessibility for faculty and student research projects to researchers worldwide.

The Philosophical Basis for Institutional Repositories

Many definitions of institutional repositories exist. Clifford Lynch, director of the Coalition for Networked Information, and leaders in SPARC have the most oft-quoted ones.¹ While some academic IRs arose from the need to combat or present an alternative

to the high cost of publishing journals and to engage in broad-based aspects of scholarly communication, government funding for research projects, publisher page charges, etc., not all institutions establish or implement an IR because of these issues, movements, and discussions. A rationale and definition could be one like that of Western Kentucky University's (WKU) TopSCHOLAR™:

A digital research repository, dedicated to scholarly research, creative activity and other full-text learning resources that merit enduring and archival value and permanent access *within a centralized* database that supports, reflects, and showcases the intellectual life of the University through easy searching and retrieval, and universal access and indexing.

The desire, if not mandate, to showcase, preserve, and freely access the intellectual and creative efforts of an academic institution is prompting many institutions to look anew at repositories as viable “containers” for accomplishing several initiatives relating to open access of scholarship and capturing permanent records of works and, at the same time, offering a unique database for students to use for primary research materials, as well as a place to showcase (although that term is somewhat pretentious) the productivity (or intellectual capital) of faculty members. If one's repository is part of a partnership or consortial arrangement, the content value increases by searching across repositories from a local site rather than searching massive indexes (aka “Googling”).

The tagging of deposits with disciplines that relate to the content, while not a perfect system, is an excellent enhancement to create a body of like-minded researchers with selected interests across repositories. Each document residing in an IR has a permanent handle, or URL, that does not change and can be readily shared with colleagues directing them to the full-text version. The ability to set up notification profiles when new content matches one's interests (similar to TOC alerts) is a value-added feature.

Not until late 2004 did IRs really gain momentum in the library world and only within the last few years have universities and colleges begun building IRs, either with open source software (and some with commercial support) like DSpace, Fedora, Greenstone, pre-prints in arXiv, or hosted services like bepress's Digital Commons. Since 2004, several key studies have now formed a basis for further data gathering and assessment. Most notably, at the University of Rochester, Susan Gibbons, Associate Dean for Public Services & Collection Development, River Campus Libraries, and Nancy Fried Foster, an anthropologist for the University of Rochester's River Campus Libraries, received a Mellon grant to study faculty research behaviors with an eye towards populating IRs and generating faculty content. The findings of Gibbons and Foster frequent discussion lists, journals, and reports concerning the values of repositories, the harvesting of content, and strategies for faculty participation.

Another Mellon Foundation Grant for the University of Michigan's Project MIRACLE (Making Institutional Repositories a Collaborative Learning Environment) and the

Council on Library and Information Resources (CLIR) studied how institutions are using IRs, best practices and policies. In 2007 Kathlin Smith from CLIR reported on a survey:

More and more academic institutions are creating institutional repositories (IRs) to manage intellectual output of their faculty, students, and staff. As such output is increasingly both produced and consumed in unmediated digital form, IRs are emerging as a critical component of the scholarly information system.²

Smith also observed that we need not policy ourselves to oblivion when first establishing an IR: “It may be more expedient to evaluate what happens after a period of time, firm up existing policies and implement new ones as needed.”³

Practicalities and Strategies

Where does one start and what does one need to consider in establishing an IR for an institution or organization? And, once established, what strategies must be addressed for the first year (usually termed a pilot project) and for the long term? The following points are key elements for introducing an IR on a campus, but, like most certainties, these can vary according to a particular group, needs, or time constraints. To paraphrase the March Hare and Mad Hatter from *Alice in Wonderland*: “Start at the beginning and when you get to the end, well ... Stop.”

- Start with a Task Force to present recommendations. The idea or directive will have come from somewhere, so get a planning group going from Day 1, the beginning.

- Develop a statement of purpose to convey what an institutional repository is and what will be placed there. What is the collection policy? Where will the IR fall within library and university priorities, and who will manage it?

Between this point and the next lies this discussion:

Have a philosophical and financial commitment from the “top down” in the university administration and the Library leadership for this collaborative effort. If the repository is hosted elsewhere, the institution’s Information Technology (IT) staff does not necessarily need to be involved but they should be informed of this movement. If library or campus IT is going to customize open source software, then obviously they must be in from the beginning. Many libraries are fortunate to have well-staffed systems departments; others must rely on the university’s staffing—and that should be a major consideration in whether to host and use one’s own server space or outsource to a ready-made platform and space. How will the system run, be serviced and supported, be customized, be enhanced? What will the response time be to problems and changes? Where is the IR priority within an IT department?

- Weigh the financial implications of hosted or local (open source software) repository. What do you get for what you are paying? Nothing is truly free. Both

architectures have associated costs. The hosted system has an annual subscription fee; the local one has staffing considerations and server space. Calculate the total cost of ownership.

If the initiative receives special funding for a year's pilot project, as many institutions do, and the IR succeeds, what plans are in place for subsequent years? In a very unscientific poll of Digital Commons users queried about funding sources, most responded that their libraries had or most likely would end up paying for the service and administering it. At one college, the provost contributed 25% of the cost. Perhaps a shared funding model whereby each college dean contributes towards the annual fee will ease the finances and allow everyone to realize the repository *is* institution-wide, not just another library database or digital library project.

Regardless of how costs are distributed, think about long-term options. A pilot IR is almost certain to succeed. Be prepared for that second year of funding!

- Acknowledge and be knowledgeable about the university's intellectual property policies and ethics policies. Create a separate copyright form for authors to deposit content in the repository. The university's lawyer or another administrator with suitable background can draft, review and also work with the service provider to craft suitable terms not only for the copyright form but for the service contract for the university if using an outside provider.

- Identify content for the launch date. Decide on some early adopters and types of content to showcase first. Start on a small scale with diverse content to showcase the value of IR. The most natural content is electronic theses, be they student honors theses or graduates students or both. Tap into web sites already within departments and find working papers, presentations, conferences on campus, publications highlighting works of faculty, staff, and students with faculty sponsors, and contact those people. Select advocates early in the process, perhaps by establishing a university-wide advisory committee, so that these researchers can offer their own works or suggest others.

- Determine how content will be uploaded. Paul Royster offers “DIFY” (do it for you) as opposed to having faculty do it themselves-- easy as it may be to upload content. For initial and even subsequent database building, do it for them! Gradually identify savvy, dependable series administrators who can shoulder certain specialties and the top-level management can experience delegation at its best.

- Market and publicize the values and benefits. Create a marketing plan, regular outlets for implementation, training, and information, and ongoing meetings with individuals, departments, councils, etc. Seize every opportunity. For librarians, create talking points about the benefits and values so they can engage in conversations as subject liaisons to the faculty.

This new role for librarians represents a huge shift when one really grasps what librarians need to be doing to support, promote and contribute to the effort. IRs represent a new dimension in collection development. Librarians are indeed building a database for students to see research (and get their own posted), for faculty to showcase or present ongoing instruction, for primary source documents to be uploaded--endless possibilities.

Get the library administration to strengthen this initiative by building "credit" into annual faculty evaluations, credit not only for soliciting content that results in deposits, but also for depositing their own content.

- Name and publicize the repository something other than an institutional repository. Early in the development of repositories on campuses, Susan Gibbons stressed this point, believing that the phrase conveys a mandate-like quality.⁴ At WKU our current, soon-to-be registered name is TopSCHOLAR™, The Research and Creative Activity Database of WKU. Other names also gravitate towards scholar, such as Scholarly Commons@, Scholar Works@, eScholarship@, eCommons@Cornell and many hosted by bepress are DigitalCommons@. UR Research is the name for the University of Rochester's and ScholarsArchive at OSU represents Oregon State University.

Be ready for challenges and opportunities. Implementation brings to fruition what was recommended, planned, marketed, and launched. The idea is reality. Strategies, while

subject to testing and modification, lay the groundwork with intensive training and initial content.

The Second Year and Beyond: Assessment of Success

So, the launch came and went. Press releases, radio interviews, brochures, training, information sessions, and advisory and management committee meetings represented time well spent in experimenting with best ways to reach people, spread responsibilities, and keep on keeping on. The experimental phase is over. Reality and the growth sets in, or the one-by-one phase. Serendipitous moments may occur, like the phone call from a professor in physical education and recreation wanting to launch an electronic journal of student research in exercise science just after you finished initial training and-- less than one year later, *The International Journal of Exercise Science* has its first issues, and a second journal is forthcoming.

Or a doctoral student from the University of Kentucky discovered two non-Shaker music pieces and two very rare printed hymnals in the library's special collections. He was told about our collections by a professor in the Music Department. He spent an afternoon in the Kentucky Library making detailed notes about for a bibliography of manuscript and printed sheet music that will be a by-product of his dissertation and funded by the American Bibliography Society.

So, how can one assess the continuance of an IR? Everyone starts with a baseline of zero, so Year 1 is a no-brainer. From zero to 500, 1000? From one community, one series to twenty? Variety in content, number of downloads over the year, top downloads—all form a basis for ongoing efforts and accomplishments. Larger research, doctoral-granting institutions have instant content with batch uploads of dissertations. Look for a group of documents already “out there” at the university.

Year 2 builds upon this groundwork and one-by-one effort grows as the IR becomes an inextricable part of the university’s scholarly landscape. The “container” swells as content increases. Calls from the Honors College to promote TopSCHOLAR™ downloads in its promotional literature, emails from the provost reminding participants to send appropriate presentations from Engaging the Spirit annual faculty conference, meetings with the Graduate Council that finally result in the first two masters theses uploaded, and always (did I mention) those one-to-one contacts.

The most recent vote and movement this spring at Harvard University Faculty of Arts and Sciences, led by Stuart Shieber, professor of computer science is certainly another example, maybe not serendipitous but deliberative, of content coming to IRs.⁵

Success *will* come, but the commitment to nurturing that outcome is relentless. Each institution is different. Commitment from administrators is essential. Once identified and defined as an integral role for the library--to build this system for the institution—librarians become the true force behind it.

Keep the connections going; generate excitement about this new pathway of discovery. As Charles E. Glassick notes, “the process, the outcomes, and especially the passion of discovery enhance the meaning of the effort and of the institution itself.”⁶ Discovery, accessibility, and permanency are the cornerstones of any institutional repository. The journey is challenging; not everyone will possess the same level of interest as you do, despite incentives. Unlike the Mad Hatter and the March Hare, once you’ve started at the beginning, there really will not be an end, for once you’ve started, you experience endless and unparalleled opportunities to be part of a growing effort in digital scholarship, scholarly communications, and opening up access to content and publishing opportunities that never existed before in this way. Digital repositories really do create information possibilities.

End Notes

¹ Clifford A. Lynch, executive director of the Coalition for Networked Information (CNI), defines an IR as “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members” in “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age,” *ARL Bimonthly Report*, no. 226 (February 2003), <http://www.arl.org/newsltr/226/ir.html>, (accessed 23 April 2008); Raym Crow, SPARC’s Senior Consultant in 2002, “... focus on scholarly communication and on changing the structure of the scholarly publishing model, we will define institutional repositories here—whatever else they might contain—as collecting,

preserving, and disseminating scholarly content,” in *The Case for Institutional*

Repositories: A SPARC Position Paper,

http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf (accessed 23 April 2008);

Rick Johnson, “Institutional Repositories Partnering with Faculty to Enhance Scholarly Communication,”

<http://www.dlib.org/dlib/november02/johnson/11johnson.html> (accessed 28 May 2008).

²“U.S. Institutional Repositories: A Census,” *CLIR issues*, no.55 (Jan./Feb.2007): 1;

information about the MIRACLE project is available at

<http://miracle.si.umich.edu/about/overview.html>.

³ *Ibid.*, 4.

⁴ These points were modified from the Institutional Repository Task Force Report at Western Kentucky University, http://digitalcommons.wku.edu/top_pres/2/.

⁵“ A Shot Heard 'Round the Academic World: Harvard FAS Mandates Open Access,” *LibraryJournal.com* (Feb. 14, 2008),

<http://www.libraryjournal.com/info/CA6532658.html> (accessed April 17, 2008)

⁶ *Scholarship Assessed* (San Francisco : Jossey-Bass, 1997), 9.