Implementing Open Access Policies Using Institutional Repositories

Ellen Finnie Duranceau and Sue Kriegsman

Implementing an open access (OA) policy offers libraries an unusually high level of challenge. Chris Armbruster, who surveyed early policy implementers says that "open access policy implementation is a tough job. Policy pioneers have faced considerable challenges in meeting their own aims and achieving recognized success."[1] But the implementation process also offers a proportionally high potential for positive payback not just to the campuses, but to the academy and the world beyond. Given this level of challenge and potential impact, libraries would do well to confer with those who have travelled further down the path, in order to maximize their chances for success. Yet not much has been written to date about policy implementation, no doubt because this task is so new to libraries.[2]

In his 2011 article, Armbruster looks at research-funder and university implementations on three campuses—Queensland University of Technology, University of Zurich, and University of Pretoria—focusing on the "infrastructure that a policy requires; the issue of capturing content…; how to provide access to the content and foster usage; and the benefits offered to authors."[5] In the U.S., open access policies are generally campus-based and are initiated by faculty but are implemented by libraries. See figures 1 and 2. Armbruster notes that policy implementation is a job that is falling to libraries and his survey reveals that librarians increasingly assume responsibility for submissions.[6] Reviewing policy implementation in the United States, it is clear that faculty-driven policies dominate the landscape and implementation therefore carries unique challenges and requirements.

OA policy implementation is a new—though natural—role for libraries, extending existing responsibilities for managing campus repositories, aggregating and curating locally created works, assigning metadata, and providing outreach on campus. While open access policy implementation is a natural fit for libraries, it is also a fundamentally different kind of implementation for libraries because it depends on the faculty for its success.[7]

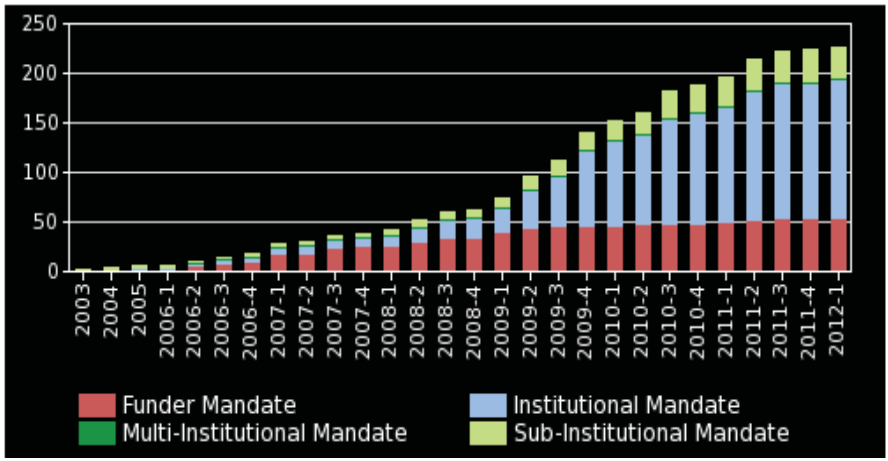

Figure 1. Worldwide, institutional open access policies are the fastest-growing type of open access policy, as is evident from the graph found at the ROARMAP: Registry of Open Access Repositories Mandatory Archiving Policies web site in the summer of 2012.[3]

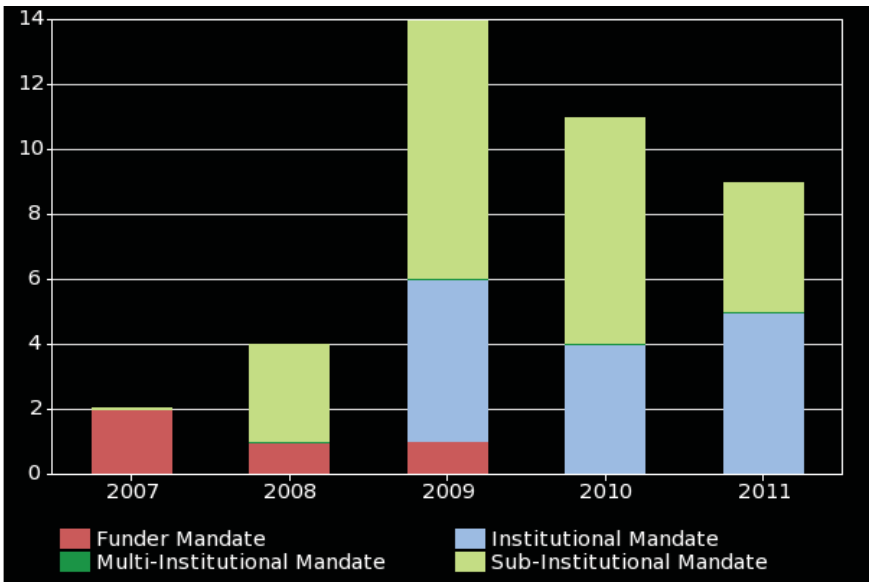

Figure 2. In the United States, where OA policy growth since 2009 has been exclusively campus-based, rather than research funder-based, institutional OA policies are a particularly important part of the scholarly communication landscape.[4]

Because of this faculty-driven quality, differences in campus cultures, faculty desires and expectations, and organizational idiosyncrasies, any single campus' perspective has only partial applicability for others. To try to make this discussion as useful as possible, then, it will be helpful to refer to the practices of multiple institutions. For this reason, an informal survey of six open access implementations, focusing on those that have had the most experience with open access to date, was carried out to determine whether any common practices are emerging in this evolving and expanding area of library services, and if so, to describe those practices. Staff at Harvard University, the Massachusetts Institute of Technology (MIT), Columbia University, Duke University, Oberlin College, and the University of Kansas were interviewed and the information obtained has been supplemented with published sources, including a short report on the policy implementation at Rollins College.[8] The eight common practices adopted by these institutions were distilled from the interviews and should be relevant to a wide audience.

While an OA policy creates an important manifestation of the will of the faculty and allows for resources for supporting open access, these emerging common practices may apply to any institution that collects articles in a repository and shares them. Thus, the common practices may apply to any institution, whether or not the institution has an open access policy in place.

Common Practice 1. Follow the Faculty

According to the sites surveyed, faculty values play a key role in motivating participation. Faculty have a common commitment to sharing their work for the greater good. For example, Duke's Director of Scholarly Communications, Kevin Smith, reported that the university's strategic emphasis on "knowledge in the service of society… really means a lot to the faculty."[9] At MIT and Harvard, the faculty's commitment to sharing is manifest in the first line of their open access policies, which states that the faculty "is committed to disseminating the fruits of its research and scholarship as widely as possible."[10]

While faculty generally agree on a desire for wide dissemination of their work, the kind of policy they are comfortable with and how they see their vision being fulfilled, and therefore what kind of policy they create, will vary. Creating and implementing a policy will depend on campus politics and governance. For example, at Columbia the open access resolutions are tailored according to the needs of each school, so that discussions with the School of Engineering centered around explicitly including conference papers in their open access resolution, whereas that type of document was not addressed in the proposal for the Lamont-Doherty Earth Observatory. On the other hand,

at Harvard, policies for each school are fundamentally the same, but the faculty of each school have made a unique commitment to an open access policy.

The Rollins experience also reflects this point. Jonathan Miller recommends that campuses "find the message that resonates with particular audiences on your campus." For Rollins faculty, that meant emphasizing "a more open system with more visibility for their own research." This goal was different from that of the provost, who "was interested in institutional reputation," or the dean of faculty, who was "interested in the idea of a stable repository of faculty publications."[11]

To implement an open access policy successfully, then, it is important to know an institution's cultures and have a clear appreciation of the wishes of the faculty. The role of the librarian, who often has the advantage of a broad perspective, is to share information and knowledge with institutional stakeholders, to absorb and synthesize diverse points of view, and ultimately, to uphold the policies the faculty have selected.

Common Practice 2. Build Trust through Outreach

Having an open access policy is dramatically important in terms of establishing the will of the faculty and providing a mechanism for addressing unnecessary limits on sharing scholarly articles, but it is not particularly meaningful unless articles are actually collected and made available under the policy. As Armbruster notes, "adopting a policy does not mean that open access will happen, it only opens up the possibility of achieving some measurable success."[18] Success comes by implementing a policy in such a way that more research and scholarship is openly available. Active outreach is critical: "in all the best-performing institutions in terms of percentage of their outputs that can be found in the repository, there is a strong, sustained advocacy program."[19]

There are a range of successful options to be considered when setting up a strong outreach and paper capture process. Effective methods include highly targeted outreach by a central scholarly communication office, student outreach, and liaison outreach, provided that the method is in keeping with the campus culture and the pace and scope of the implementation are supported by the faculty. The key is to stay in touch with the campus context.

The University of Kansas launched its implementation by targeted outreach, working initially with known supporters. Ada Emmett, scholarly communications librarian at the University's Center for Digital Scholarship, reports that "we proceeded carefully—but with a balance of tiny staff and wanting to build friends and early adopters first and not have a full-court press going forward,

Policy Types

Faculty need to create their own OA policy, but librarians can provide details about the relative advantages of the different policy types. The first faculty in the United States to adopt a policy that granted permission (a license) to post articles online was Harvard University's Faculty of Arts and Sciences, which voted unanimously in 2008 to give the university a nonexclusive license to their scholarly articles. This permission-based policy approach was followed by various other Harvard faculties and then by the Stanford School of Education, MIT, the University of Kansas, Oberlin College, Duke University, Rollins College , the University of Hawaii at Manoa, Columbia's Lamont-Doherty Earth Observatory, Princeton University, Bucknell University, Utah State University, and the University of California San Francisco, among others.

Permission-based policies have more legal power than other kinds of policies and resolutions because they extend a license to the university for sharing the articles that predates a publisher contract, so the publisher's contract is subject to the prior license. According to legal scholar Eric Priest, the permission-based policy "is a far more powerful open access tool… than its deposit-mandate cousin" primarily because a permission-based policy "resets the default in copyright law."[12] Permission-based policies are opt-out policies. They invoke a license automatically from the time a scholarly article is written and they apply unless the author requests a waiver. Opt-out policies have been shown to create higher participation rates across many different domains. Stuart Shieber, the initiator of Harvard's policies, reports that "in many areas (organ donation, 401k savings), participation tends to be much higher with opt-out… systems, and that holds for rights retention as well."[13]

Other types of policies represent support for open access or at least some level of commitment to deposit a scholarly article when a publisher's policies allow. For example, Boston University's (BU) resolution calls for building an open access repository and promoting open access in routine operations.[14] Resolution-style policies are significant because they represent the will of the faculty and have the potential to be powerful forces for change. BU's policy calls for changes in promotion and tenure processes.

Resolution-style policies create organizational commitment and can justify diverting resources to support open access. However, policies without a license or permission element do not create a legal right to post articles, as do permission-based policies. Similarly, policies that call for action only if that action aligns with existing publishers' policies do not have the same impact as a permission-based policy. Resolution-style policies provide a significant catalyst for change, but in and of themselves they do not alter the legal framework for authors as do permission-based policies. For this reason, Alma Swan refers to resolution-style policies as weak, and says they compromise the public interest.[15]

Nevertheless, libraries with no policy or a weak policy can achieve more open access to campus research and scholarship by including author-rights language in library content licenses or using the many friendly publisher policies, particularly if the faculty has given voice to a desire to make their work openly available.[16]

where we wouldn't have the staff to manage if successful or the good will currency if troubles came forward." Their focus was to "build up a cadre of adopters" in a way that was "organic rather than systematic." One way they built trust and gained early adopters was asking distinguished professors to participate.[20]

Columbia has also had success with targeted outreach. There, the outreach has gone beyond the two units with a formal policy, the Lamont Observatory and the Libraries. Robert Hilliker, Columbia's Digital Repository Manager, reports that "contacting people who've been in the news gives the highest rate of return."[21] They have also found that it is productive to focus on interdisciplinary centers. Similarly, Oberlin began by seeding the IR through focused outreach and reaching out to the faculty coordinating their policy and asking them for resumes to review and requesting manuscripts as appropriate.

Harvard did targeted outreach to a few departments to prepopulate the repository prior to making the content accessible to the public. This strategy gave the repository a robust look on day one but also gave the staff time to troubleshoot and debug as the first deposits were received. Once the repository was publicly available, there was additional outreach at department meetings. As the repository matures, a new round of department outreach has started.

In order to prepopulate and test deposits during the early stages of their repositories, both Kansas and Oberlin reviewed the *curriculum vitae* (CV) of each faculty member to identify eligible papers. In lieu of a formal policy, Columbia carries out an ongoing review of CVs to identify papers for inclusion in the IR. At the Harvard Graduate School of Education, library staff regularly review faculty web sites and CVs to discover current publications that are appropriate for deposit in the repository.

MIT has taken a systematic approach, working department by department, reaching out to each department and author once per year. Library staff first meet with the department chair, if possible, and then try to arrange to speak with the faculty at one of the department's regularly scheduled events. MIT also aims to have department chairs send an email encouraging faculty to respond to a call for papers before sending requests to faculty for papers. They have found that results are generally better when these steps are taken.

As part of the outreach activities, another key practice is informing authors of publisher policies, even when a permission-based policy makes this unnecessary for posting. Authors often want to know "What does my publisher think of the policy?" This work needs to be carried on through many channels: one-on-one visits with authors and department heads, presentations at departmental faculty meetings, and contact with key committees.

Acquiring Papers: Which Papers and from Where?

Provided it is permissible for the documents to be copied, identifying sources for papers is one of the key questions that must be addressed and resolved, including where papers will be obtained and whether to take advantage of appropriate copies of papers that are already available on the web through various web sites such as the campus web domain, subject repositories, and publisher web sites.

MIT requests papers from authors only when they cannot be obtained any other way. Their policy states that MIT will "develop and monitor a plan for a service or mechanism that would render compliance with the policy as convenient for the faculty as possible."[22] This means that if a lawful copy can be obtained from a publisher, a repository, or elsewhere online, rather than ask authors for papers, they use other sources for the paper. Duke also obtains papers from publisher web sites, as permissible and appropriate.

At Columbia, authors who deposit in open access repositories such as PubMedCentral, arXiv, or SSRN, are in alignment with the policy, and their papers are not collected or deposited in Columbia's IR. At MIT, in contrast, papers available through other repositories are collected for deposit, as permissible. This practice is driven by the faculty's view that it is important not to rely on external sources to make MIT scholarship available, a point that was later reinforced by the threat of the Research Works Act, which would have made illegal the NIH Public Access Policy's requirement that government-sponsored research articles appear in PubMedCentral.[23] Harvard also has decided not to rely on external sources to collect faculty papers.[24]

Common Practice 3. Repurpose Existing Staff

Implementing an open access policy via a repository has been primarily managed by repurposing staff and reprioritizing the duties of existing staff rather than adding staff, although many sites have redefined or created one position largely devoted to implementation. In instances such as Harvard, where a repository did not already exist, staff were added to introduce and manage the repository. Funds for implementing an open access policy are often soft money, including grant funds, budget transfers, gifts, or endowment income. Permanent staff may be supplemented by interns, temporary workers, and students. One-time funds may be available for new positions needed to support open access activities.

Open access policy implementation is generally managed by someone who has at least 25 percent of his responsibilities assigned to scholarly communication. At Oberlin, an associate library director was given the title of "schol-

arly communication officer" and assigned duties that included responding to questions about the open access policy and providing assistance in interpreting publisher policies. Princeton first appointed a librarian to serve as "acting scholarly communication librarian" responsible for a team of repository librarians. After a period of time, the position was made permanent. The implementation of Columbia's open access policy was directly associated with repository services. At MIT, however, the scholarly publishing librarian oversees OA implementation and works closely with the repository management staff and with liaison subject specialists through a policy-related outreach committee. Some universities have a centralized scholarly communication department where the repository and policy support are jointly managed by the same team, as is the case at Harvard. At least one site reported finding that communication could be complex when not tasked to a single department or a close-working interdepartmental team.

Staffing levels vary with campus scale and other factors, but in general, sites reported that two or three individuals were heavily involved.

- At Kansas, 1.5 staff FTE, and an additional 2 or 3 students, with each student working 10 hours a week, manage the entire process.
- At MIT, nearly 1 FTE (parts of two librarian positions) work heavily on the implementation, with support from a ten-hour per week temporary staff member who assists with the deposit of papers and 0.5 FTE support staff from acquisitions who helps to acquire and deposit papers. In addition, library liaisons each spend about 5 hours per department per year requesting papers and doing follow-up; metadata staff offer a total of between fifteen and thirty hours per week cataloging the papers; and the repository manager devotes a significant portion of his time to OA implementation issues, including problem-solving and gathering and sharing usage data. In addition, one librarian spends approximately four hours per month on database management.
- At Harvard, 2 FTE librarians are responsible for policy implementation and management, 0.5 of a faculty FTE oversees the scholarly communication office, and 1 FTE software developer is devoted to supporting the project in DSpace and other departmental activities.
- Columbia, MIT, and Harvard use library liaisons to reach out to faculty to request papers; pilot projects at Kansas and Duke involve liaisons, and other sites plan to follow that example. At Oberlin, the liaison's role is seen as a "natural fit" for outreach to faculty according to Director of Libraries Ray English. Following a pilot program with six departments from November 2009 to March 2010, MIT officially implemented its liaison program in 2010. During the pilot, liaison librarians at MIT found

> that engaging in outreach enhanced their relationships and provided opportunities for discussions with faculty, some of whom had never before availed themselves of the liaison's expertise. For example, MIT physics librarian Mathew Willmott reported that the experiment offered a "good opportunity for outreach about scholarly publishing." The liaisons who participated in the pilot concluded that the effort was a "good way to get to know faculty publications… [it] created opportunities for positive interaction… [and] cultivated relationships with faculty." Being put "in the publishing workflow" was seen as a boon to their relationships with faculty, a sentiment often echoed by Harvard's liaisons as well.[25]

Based on this success, and in the context of a reorganization and related redefinition of liaison roles, the MIT Libraries officially added outreach about the open access policy to the job description for liaisons:

> Provide support for strategic scholarly publishing activities and provide services, such as recruitment of faculty authored research materials and promotion of repository-based services.[26]

MIT liaisons are provided with lists of faculty-authored articles that they mail to faculty, requesting manuscripts for the articles. Prior to contacting faculty, liaisons try to meet with the department head and a member of the Office of Scholarly Publishing and Licensing to discuss the best approach for outreach to each department. For example, one department requested that messages be sent in the last week of January.

At Harvard, seven students, each working from five to ten hours per week, recruit papers, create metadata, and stage the papers for deposit into the repository. Each school has a designated open access liaison who maintains contact with the faculty. Although most of the liaisons are librarians, other staff, such as those associated with grants or research offices, also serve in this capacity. They reach out directly to the faculty while students complete the deposit process. The liaisons have successfully created an informal resource community and have established customized relationships with the faculty in each school while also developing expertise about Harvard's open access policies and open access in general.

Common Practice 4. Make It Easy

Columbia's Rebecca Kennison could be speaking for all the sites when she says that faculty are "fine with their stuff going into the repository, but they don't want to have to *do* anything (or at least as little as possible) to make that happen." As a result "all [of Columbia's] processes are designed to make it as easy

as possible for [faculty] to deposit."[27] In keeping with that principle, Columbia, like MIT and Harvard, uses a self-deposit form that requests minimal metadata. MIT asks only for the author's name and the title or the Digital Object Identifier (DOI) of the item; they accept any version of the paper and convert it to PDF as needed. Similarly, Oberlin's Ray English explains that they "do all we can to make it easy for faculty," including, like MIT and Harvard, allowing faculty to submit papers via email rather than requiring them to use the submission form.[28]

Since faculty desire an easy process for contributing papers, it makes sense to insert the repository process into their existing workflow as seamlessly as possible. More than one site harvests content from web sites to which faculty already make deposits, such as open access subject repositories, personal profile sites, and local activity reports. In 2010 Harvard began including an "upload to DASH" option in the Faculty of Arts and Sciences annual faculty reporting tool. A follow-up email asking for specific papers identified in the activity report that had not already been uploaded resulted in a windfall of about four hundred new papers for the repository.

MIT has set up automatic ingest with BioMed Central and SpringerOPEN so that articles are automatically deposited via the Simple Web-service Offering Repository Deposit (SWORD) protocol, and is working with Hindawi in the hope of setting up a similar system.[29] This automated deposit streamlines MIT's deposit time dramatically and results in all MIT-authored papers being deposited, with metadata, directly into DSpace@MIT. The last step is identifying MIT authors using name authority control and department information that is not available to BioMed Central.

When permissible, Harvard and MIT harvest open access articles from subject repositories such as arXiv and PubMed Central. Since the faculty already deposit to these repositories, it makes sense to take advantage of the repositories' open APIs. At Harvard, affiliates listed in the subject repository are matched against the staff directory. Papers by Harvard authors are then downloaded and prepared for deposit. Although a human component is needed to check the names and articles, the process is fairly quick and streamlined.

Common Practice 5. Buy or Build an Infrastructure

While having the faculty's trust and an understanding of their vision are essential to successfully implementing an open access policy, the infrastructure for obtaining and depositing papers is also necessary, particularly if a systematic workflow and the means of assessing progress are desired. As with all aspects of open access policy implementation, the practice must fit the campus climate,

needs, and resources. As Armbruster notes, "implementation is a job that requires developing a match between the institution and the infrastructure."[30]

While an institutional policy on open access is not essential, having an institutional repository (IR) to collect and archive campus scholarship is highly desirable. At Duke University and at MIT the repositories were up and running for years prior to the adoption of an open access policy by the faculty. The repository at Harvard was created as a direct result of the first OA policy, as will be the case at Princeton. If the policy precedes the IR, there can be a delay in policy implementation while the repository is built, which could result in lost momentum. However, Rollins College's Miller, whose faculty passed a permission-based policy similar to Harvard's, points to additional reasons for recommending that the repository be built while the policy is still under consideration, rather than after it is passed. He reflects that while simultaneous policy and repository development "can be expensive in either time or money and could result in a successful implementation of the repository," he concludes that no policy "is not necessarily a bad thing. Building the repository at the same time as you press for the policy means that faculty will be able to see practical examples of how their works will be archived and accessed. This can reassure faculty in disciplines that are not already making extensive use of disciplinary repositories."[17] Rollins launched a repository simultaneously with the passage of a policy in 2010.

The sites surveyed for this report have all grappled with developing some kind of technical infrastructure beyond the repository to support implementation, and all leverage existing publication information from citation databases and in some cases, from faculty CVs. While several sites use spreadsheets to manage the data, others, including MIT, have created databases for this purpose.

Duke, Oberlin, and MIT use Web of Science® to identify citations for papers published at their institutions as a basis for paper acquisition, including sending messages to authors requesting their manuscripts. MIT feeds citation information into a locally designed relational database, while Oberlin uses a spreadsheet. MIT and Oberlin use reference management tools to manipulate the references identified from abstracting and indexing services.

At least two sites are seriously examining whether a commercial tool such as Symplectic's Elements should be purchased to take the place of the data acquisition step as well as other aspects of workflow and services to authors.[31] All sites are working to leverage legacy faculty data systems to obtain publication information, to reduce redundancy, and to make data flows as efficient as possible.

Most sites use the infrastructure they have built, whether spreadsheets or databases, as a basis for building email reminders to faculty about papers that could be deposited under the open access policy.

Duke harvests citations and articles from databases and then determines whether those versions can be added to the repository or, if not, whether they can acquire a copy from another source. They send authors an email alerting them to the fact that their article is now in DukeSpace. Authors are then given the choice to opt-out or add items. Duke takes this activity one step further, working with a commercial company, Symplectic, to harvest citations from published sources and integrate data from legacy faculty data systems. The collected information is sent automatically to the faculty profile system used to populate the faculty web pages. One advantage to this choice of working with a vendor is that the vendor has built software to deduplicate and disambiguate the harvested materials and will automate many of the workflow processes.

None of these systems can provide a complete list of faculty scholarship. As Alma Swan notes, "checking the repository content against what is recorded by literature indexing services gives only an approximation of how complete the repository's content is [since] there is no indexing service that covers 100 percent of the literature."[32]

While each site has different workflows and uses different tools, MIT's system is well developed and it illustrates many of the issues and needs involved with an implementation.[33] See "MIT's Workflow System."

Synergy with Other Campus Publication Tracking Systems and Services

Once the library is engaged in implementing an open access policy, inevitably questions arise about how to integrate open access with other campus activities involving publication reporting and related data. Oberlin's College of Arts & Sciences and the Conservatory of Music are considering reinstating a process for creating an annual summary of activities, including publications. According to Library Director Ray English, the challenge is "how to combine the repository function, the open access policy, annual activities, and the listing of publications and performances." In a statement that could apply to any of the sites, he comments that it is "difficult to bring these all together but we are trying to develop a synergy between the… needs."[34]

Miller reports that at Rollins, data is drawn from "annual reports each faculty member submits to the dean."[35] This method has the strong advantage of

providing very complete information but also builds in a potentially lengthy delay prior to learning about a given publication, making the acquisition of papers more challenging, since faculty are less likely to be able to find the final draft of their paper as time goes on.

Sites are still working towards seamlessly meshing data from indexing services with rich metadata gleaned from CVs updated on authors' web sites and annual reporting mechanisms, which tend to be the most complete sources.

To move in this direction, Columbia's Rebecca Kennison points out that it can be productive to "cultivate our relationship with the Office of Research." This kind of relationship can help with the coordination of data that is needed and used in various ways by diverse campus groups and also raises the library's visibility and associates repository collection-building as a "valued part of the research enterprise of the university."[36]

Common Practice 6. Add Value

Metadata

Sites usually provide a full citation and pointers to the final published version from their repositories. Most publishers request this, and it is a valuable service for readers as well. Some sites are also doing name authority control which is an important, if not essential, foundation for access and the basis for other repository services, such as author-specific usage data. In addition, librarians are keeping an eye on the Open Researcher & Contributor ID (ORCID) project which aims to disambiguate author names in scholarly communication by utilizing a central registry for author names and works.[37]

Beyond that, metadata choices vary, but the surveyed sites are finding, for the most part, that their work on metadata for open access policy implementation positions them to be relevant to their campuses in other ways. The information is in high demand for many reasons, including for support of administrative, communication, and public relations goals. It can also play a central role in a wide range of services and tools that profile faculty research or assist researchers in finding collaborators.

MIT uses the Scholarly Works Application Profile (SWAP) for the metadata schema and records which deposit policy applies to a paper, for example the publisher's policy for papers prior to the adoption of the MIT Faculty Open Access Policy, or the Open Access Policy that provides MIT a license to share the work openly. MIT also describes the type of content (journal article versus conference paper) and the version of the paper and its peer-review status, as well as an essential piece of metadata—the terms of use.[38] Papers submitted

MIT's Workflow System

Since MIT did not have and does not maintain a central database of faculty publications, the first task in developing a workflow system to support the faculty's open access policy was gathering citations to MIT-authored papers by accessing several licensed databases. Using a Microsoft Access database and scripts devised by MIT Physics Librarian Mathew Willmott, citations identified through alerting services in each database are imported into a citation management tool.

After normalizing the data and checking for duplicates, the citations are exported from the citation management tool and imported into a custom-built relational database known as the OA Workflow Database. This database was developed as a prototype to provide a basis for acquiring papers and, as a byproduct of this work, to identify system requirements. See figure 3.

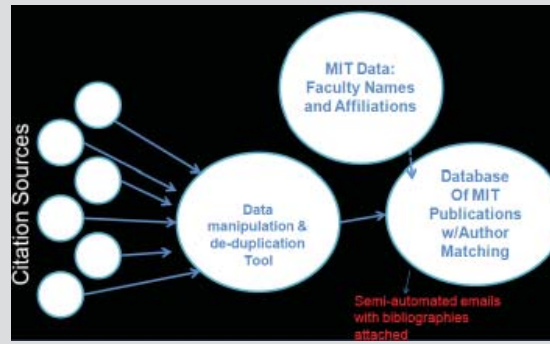

Figure 3: MIT workflow for a database supporting open access policy implementation workflow

Information obtained from the university's Human Resources Department is imported into this database and a script matches faculty members with their papers. This step is needed because unfortunately, searching by affiliation in various citation databases does not consistently pair specific author names with the MIT address or affiliation.

The system includes workflow controls to mark progress on obtaining each paper; notes to track interactions with faculty members; and an automatically generated email template to contact faculty about outstanding papers. See figures 4 and 5.

Tables of publishers and journals track various policies with respect to author rights, harvesting, and version requirements. See figure 6. Queries generate statistics on facets of the project and allow for a thorough assessment of the process.

Where there is permission to do so, MIT obtains papers from publishers' sites using a "Find Full Text" service to download papers. Under the SWORD protocol, papers are automatically delivered/deposited into the repository, harvested from BioMed Central, and papers not obtained as a result of that process are downloaded manually from publisher (or repository) sites.

For papers not obtained from publishers, MIT requests

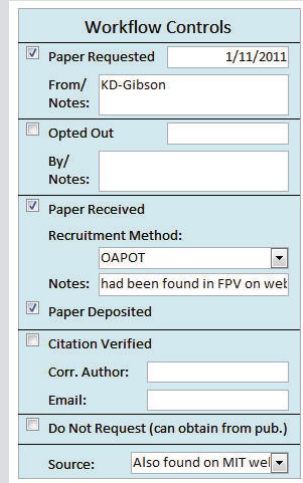

Figure 4. MIT's Open Access Workflow Database Workflow Controls

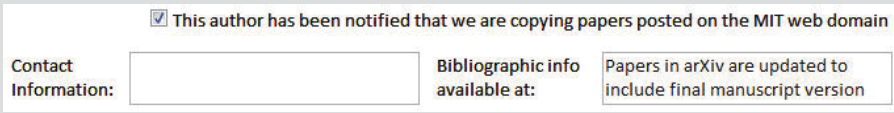

Figure 5. Communication records in MIT's Open Access Workflow Database

that authors submit their final prepublication manuscript. The OA Workflow Database supports this process: the email template generates a message containing a list of papers being sought from authors, using a script that removes articles from the authors' list that can be obtained from publisher, have already been received, or which have been opted-out. See figure 7. The database also offers the capability to label papers that authors can supply in final published form.

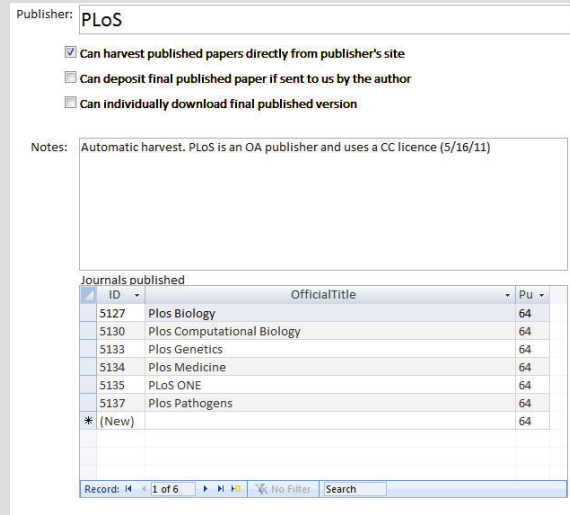

Figure 6. Publisher table in MIT's Open Access Workflow Database

Messages can be customized based on information stored about publisher policies, indicating, for example, whether authors may provide the final published version of an article. Liaisons may choose to customize or personalize the messages further prior to sending them.

Figure 7. Sample email message generated by MIT's Open Access Workflow Database

under the MIT Faculty Open Access Policy are made available under a CC-BY-NC-SA license, as it was very important to the MIT faculty who established the policy that their work be available under an open, standard license.[39]

Peter Suber supports this view and laments that more open access policies that target self-archiving have not been accompanied by terms that allow full reuse, as under a Creative Commons Attribution (CC-BY) license, what he calls "libre" open access. He says "there are more than thirty university OA policies based on the Harvard [permission-based policy] model. All of them are at least potentially libre, by granting the institution enough rights to authorize libre access through the repository. But whether they are actually libre depends on what the repository actually authorizes."[40] He points to the University of Liege and the library faculty at the University of Oregon where articles are available under a CC license, allowing for full libre open access. Suber assesses Harvard's unique license as "not CC-BY, but it's way beyond fair use."[41]

Repository-related Services

Armbruster's survey of policy pioneers outside the U.S. finds that, at a minimum, authors will need mediated deposit if the implementation is to succeed. "In the interest of building an open access collection and maintaining momentum, it then becomes necessary to archive on behalf of the authors. Self-archiving is replaced by a system best described as assisted deposit or mediated deposit."[42] The informal survey reported on in this chapter confirms that U.S. institutions are likewise performing assisted or mediated deposit.

As part of the deposit process, both MIT and Harvard reformat papers submitted as Word documents, or any other format, to PDF. Small steps such as this can simplify the process for authors and lower barriers to contribution, since the details are handled by the staff handling the actual deposit.

Harvard encourages authors to sign an "assistance authorization" form (http://osc.hul.harvard.edu/dash/proxy) empowering the Office for Scholarly Communication (OSC) to make deposit decisions on behalf of the author, and requires such authorization for any deposits mediated by the OSC. The deposit decisions include selecting the license situation for the article, since the repository offers three options from which to choose: (i) The article falls under the Open Access Policy and the author has not obtained a waiver for the article, in which case the depositor (or the OSC if authorized) agrees to an Open Access Policy (OAP) Author Agreement; (ii) The article is not subject to the Open Access Policy or a waiver was obtained, in which case the depositor (or, again, the OSC on the depositors behalf) agrees to a Limited Author Agreement (LAA) that provides a limited license to distribute the article from the repository perhaps subject to an embargo; (iii) The depositor does not have rights to allow

distribution, in which case the article is deposited dark.[43] Signing the authorization can relieve authors of making these kinds of decisions, allowing the OSC or others mediating deposits on their behalf to make the decisions for them. However, not all sites find an author's signed authorization necessary to allow for mediated deposits, reflecting again how local culture and interpretations can make a significant difference in implementations, even when operating under identical policy language.

Harvard and MIT both provide repository-level statistics to the public, including a list of most heavily downloaded papers. Through local software development to enhance standard DSpace capabilities, Harvard also provides the number of article downloads, number of visitors, and number of articles deposited. The information can be viewed for the last seven days, the last thirty days, or cumulatively, and statistics are available for the whole repository or for individual schools. Individual authors also have the ability to login to the system and view statistics about their own content. A heat map indicates the countries from which the downloads occurred. Harvard sends an automated monthly email to faculty with statistics about their articles and a reminder that more details are available at the repository site. The response from the faculty regarding access to their own statistics has been positive.

Developing services in relation to the repository can absorb as little or as much time as a site allows or allocates. There are endless enhancements and features, such as "@mire" for DSpace, that can be developed or added by purchasing repository services.[44] Harvard has decided to not focus on the repository interface because most visitors to the site never see the front page of the repository. Sixty percent of the traffic comes from Google or other search sites, 20 percent from direct referring sites, and only about 20 percent comes through the front door.

There are various ways to market the repository and additional enticements that can be used to keep people coming back to the repository and keeping it fresh in their minds. MIT created videos of faculty talking about the open access policy and its significance to the institution and for research. Harvard identifies articles that may have general public interest; writes summaries of them; and posts the summaries on the front page of the repository, the department Facebook page, and Twitter account. MIT blogs about articles in the repository related to faculty authors in the news. These methods can encourage visits and deposits.

Repository data can also help keep faculty profile pages or activity reports up to date. At Duke, after citations from published sources as well as from campus

legacy systems have been harvested, faculty will be invited to make corrections to the data. Based on information in SHERPA/RoMEO, faculty will be notified which articles are eligible to be deposited into the OA repository. Once an article is deposited in the repository, the link will be added to the faculty profile page. Harvard is currently working toward prepopulating annual faculty activity reports with citations from the current year of publications that are deposited in the open access repository.

Common Practice 7. Engage with Publishers

One significant advantage of institutional OA policies is that, either by way of a legal license to the work or because the policy represents the will of the faculty, they allow the institution to work with publishers on authors' behalf. MIT Professor Richard Holton commented in a recent article about the implementation of MIT's policy: "one of the premises of the MIT Faculty Open Access Policy was that it would make it possible for MIT to be at the table for discussions, rather than leaving each MIT faculty author responsible for negotiating their author rights alone."[45] Publisher engagement is needed even when a permission-based policy provides a legal framework for deposit. For example, implementers have had to grapple with sometimes confusing and highly ambiguous situations when the publisher attempts to negate open access policies through their author terms.

Regardless of such publisher pushback, the rate of waivers (opt-outs), remains low. At Harvard, approximately 5 percent of articles in the Faculty of Arts and Sciences have had waivers issued, at MIT, 4 percent. Further, some publishers who require an opt-out nevertheless allow posting of the final article later, including The *Proceedings of the National Academy of Sciences,* which requires a delay of six months.

None of the sites that were surveyed agree to automatic embargoes requested by publishers.

Common Practice 8. Share Information and Keep Learning

Some barriers to success are not within the control of libraries implementing open access policies. Barriers encountered by organizations outside the U.S. and identified by Armbruster sound remarkably familiar to early U.S. implementers, including limits on faculty time and faculty's reluctance to share versions other than the final, published version.[46] MIT and Harvard have found that limited time, version confusion, and concern about publisher policies

hinder the deposit process. Automating the deposit process, enabling dual-repository deposit, raising awareness, and ongoing cultural change may begin to address these issues in the future.

This chapter offers an overview of the steps early implementers have taken to carry out the will of the faculty and support them. Even with similar policy language, campus implementations vary in order to adhere to the faculty's wishes. Based on the experiences of the early U.S. implementers, it is evident that sharing information has made the path easier. Such sharing has been facilitated by the birth of the Coalition of Open Access Policy Institutions (COAPI). COAPI was formed by the University of Kansas in August 2011 with twenty-two participants, to "collaborate and share implementation strategies and advocate on a national level for institutions with open access policies."[47] Sharing information is a critical role for the organization, as so many libraries without any previous experience are attempting to ramp up with policy implementations. In its first year of operation, COAPI expanded to more than forty-five institutional members, a testament to the growing body of institutions with policies and the importance of this trend in libraries to support faculty and, more broadly, to support the evolution of scholarly publishing.

The growth of institutional open access policies in the United States suggests that the work libraries are doing is a highly relevant and important dimension of the broader goals to help faculty redress imbalances in the scholarly publishing system and share their work as widely as possible, so that it can have the greatest possible impact. At the heart of these policies lies the institutional repository, whose essential function is allowing libraries to collect, preserve, and make faculty research and scholarship openly available.

The implementation of an institutional open access policy differs from other implementations with which librarians have been engaged because librarans must follow the faculty's lead, and the implementation will succeed only if the faculty are involved and participate. This is a difference that librarians can embrace and celebrate as a manifestation of their central, trusted role. As Oberlin's Ray English notes, "every time we make a piece of research openly available, it's a gain."[48]

Acknowledgements

The authors extend their thanks to the following individuals who so generously shared their current practices or helped summarize those of others. Any errors are completely our own.

- Ada Emmett, Scholarly Communications Librarian, The University of Kansas

- Ray English, Director of Libraries, Oberlin College
- Rob Hilliker, Digital Repository Manager, Columbia University
- Rebecca Kennison, Director, Center for Digital Research and Scholarship, Columbia University
- Anne Langley, Head Librarian, Science and Technology Libraries, Princeton University
- Paolo Mangiafico, Director of Digital Information Strategy, Duke University Libraries, Duke University
- Stuart Shieber, Director, Office for Scholarly Communication and James O. Welch, Jr. and Virginia B. Welch Professor of Computer Science, Harvard
- Kevin Smith, Scholarly Communications Officer, Duke University

Reference Notes

1. Chris Armbruster, "Open Access Policy Implementation: First Results Compared," *Learned Publishing* 24, no. 2 (2011): 1, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1927775 (accessed September 3, 2012).
2. Armbruster reports that prior to his 2011 study "no substantive study of policy generation or implementation" had been published.
3. Registry of Open Access Repositories Mandatory Archiving Policies (ROARMAP) web site (2012). ROARMap, http://roarmap.eprints.org (accessed August 14, 2012).

   ROARMAP is powered byEPrints 3 (http://eprints.org/software/) which is developed by the School of Electronics and Computer Science (www.ecs.soton.ac.uk/) at the University of Southampton. More information and software credits. (http://roarmap.eprints.org/eprints/). Copyright © 2012 University of Southampton. This work is licensed under a Creative Commons Attribution 2.0 UK: England & Wales License (http://creativecommons.org/licenses/by/2.0/uk/deed.en_GB).
4. Constructed from data at the ROARMAP site (accessed February 28, 2012).
5. Armbruster, 1.
6. Ibid., 7.
7. Ray English, Personal communication, February 2012. Throughout this chapter, quotes from the survey respondents represent personal communications based on email and phone interviews conducted in February 2012 with Ada Emmett, Scholarly Communications Librarian, The University of Kansas; Ray English, Director of Libraries, Oberlin; Rebecca Kennison, Director, Center for Digital Research and Scholarship, Columbia University; Rob Hilliker, Digital Repository Manager, Columbia University; Stuart Shieber, Director, Harvard Office for Scholarly Communication and Professor of Computer Science; Anne Langley, Head Librarian, Science and Technology Libraries, Princeton Univer-

sity; Kevin Smith, Scholarly Communications Officer, Duke University, and Paolo Mangiafico, Director of Digital Information Strategy, Duke University Libraries, Duke University.

8. Jonathan Miller, "Open Access and Liberal Arts Colleges: Looking beyond Research Institutions," *College & Research Libraries News,* 72, no. 1, (2011): 16–30, http://crln.acrl.org/content/72/1/16.full.pdf+html (accessed September 3, 2012).
9. Kevin Smith, Personal communication, February 2012.
10. MIT Faculty Open Access Policy, http://libraries.mit.edu/sites/scholarly/mit-open-access/open-access-at-mit/mit-open-access-policy/ (accessed September 3, 2012); Harvard Faculty of Arts and Sciences Open Access Policy, http://osc.hul.harvard.edu/hfaspolicy (accessed September 3, 2012).
11. Miller, 18.
12. Eric Priest, "Copyright and the Harvard Open Access Mandate," *Northwestern Journal of Technology and Intellectual Property* 10, no. 7 (Oct. 2012): 377–440, http://scholarlycommons.law.northwestern.edu/njtip/vol10/iss7/1/ (accessed November 19, 2012). Full discussion of the copyright law in relation to permission-based open access policies.
13. Stuart Shieber, "Clarifying the Harvard Policies: A Response," *The Occasional Pamphlet on Scholarly Communication* (December 2, 2011), http://blogs.law.harvard.edu/pamphlet/2011/12/02/clarifying-the-harvard-policies-a-response/ (accessed September 3, 2012).
14. Boston University; University Council Committee on Scholarly Activities and Libraries; Faculty Council Committee on Research Activities, Libraries, and Support Services; *Scholarship, Libraries, and Open Access Archiving Initiative,* www.bu.edu/av/today/slideshows-and-audio/images/OpenAccessInitiative.pdf (accessed September 3, 2012).
15. Alma Swan, *Policy Guidelines for the Development and Promotion of Open Access,* (Paris, UNESCO, 2012), 49, unesdoc.unesco.org/images/0021/002158/215863e.pdf (accessed September 3, 2012).
16. For model author rights language for library content licenses see Author Rights Model License Language, http://authorrights.wordpress.com/ (accessed September 3, 2012). The concept is discussed in more depth in Ellen F. Duranceau and Ivy Anderson, "Author-Rights Language in Library Content Licenses," *Research Library Issues; A Bimonthly Report from ARL, CNI, and SPARC,* no. 263, (April 2009): 33–37, www.arl.org/bm~doc/rli-263-author-rights.pdf (accessed September 3, 2012).
17. Miller, 18.
18. Armbruster, 5.
19. Swan, 50.
20. Ada Emmett, Personal communication, January 23, 2012.
21. Robert Hilliker, Personal communication, February 2012.
22. MIT Faculty Open Access Policy.

23. Research Works Act, H.R. 3699, 112th Cong. (Dec. 16, 2011), www.gpo.gov/fdsys/pkg/BILLS-112hr3699ih/pdf/BILLS-112hr3699ih.pdf (accessed September 13, 2012).
24. Harvard University, "Overview: DRS & Delivery Systems," http://hul.harvard.edu/ois/systems/drs/ (accessed September 18, 2012). The Digital Repository Service "is the centerpiece of HL's digital preservation program," http://hul.harvard.edu/ois/digpres/repository.html (accessed September 18, 2012). At this time Digital Access to Scholarship at Harvard (DASH), a DSpace installation, is not considered a preservation repository because it does not include preservation services such as format migration. However the university is committed to preserving DASH content in its Digital Repository Service (DRS).
25. Ellen F. Duranceau, Mathew Willmott, and Michelle Baildon, *Inventing the Future of Scholarly Communication: Supporting the MIT Faculty Open Access Policy,* MIT Libraries All Staff Presentation, April 14–15, 2010.
26. MIT Libraries' internal document, liaison job descriptions.
27. Rebecca Kennison, Personal communication, February 2012.
28. Ray English, Personal communication, February 2012.
29. Ellen F. Duranceau and Richard Rodgers, "Automated IR Deposit via the SWORD Protocol: An MIT/BioMed Central Experiment." *Serials*, 23, no. 3 (2010): 212–14, http://uksg.metapress.com/content/l437x1631052407r/fulltext.pdf (accessed September 3, 2012).
30. Armbruster, 5.
31. Symplectic, www.symplectic.co.uk/ (accessed October 1, 2012).
32. Swan, 49.
33. Information on MIT's workflow system was adapted and expanded from a poster presented at the Berlin 9 conference by Mathew Willmott and Ellen Duranceau, "Implementing an Open Access Policy: Building the Workflow to Identify and Gather Content," Bethesda, Maryland, Nov. 9–10, 2011.
34. English, Personal communication, February 2012.
35. Miller, 18.
36. Kennison, Personal communication, February 2012.
37. ORCID: Open Researcher and Contributor ID. http://about.orcid.org/ (accessed September 3, 2012).
38. JISC, "Scholarly Works Application Protocol," Digital Repository Wiki, www.ukoln.ac.uk/repositories/digirep/index/SWAP (accessed September 3, 2012).
39. Creative Commons AttributionNonCommercialShareAlike 3.0 Unported (CC BY-NC-SA 3.0 License, http://creativecommons.org/licenses/by-nc-sa/3.0/ (accessed September 3, 2012). This license allows users to copy, distribute, transmit, and adapt the work. They must provide attribution for the work and may not use it for commercial purposes.
40. Peter Suber, "The Rise of Libre Open Access," *SPARC Open Access Newsletter*, no.164**,** June 2, 2012. www.earlham.edu/~peters/fos/newsletter/06-02-12.htm 6/2/12 (accessed September 3, 2012).

41. Suber.
42. Armbruster, 7.
43. Harvard University Library. Office for Scholarly Communication. Assistance Authorization, http://osc.hul.harvard.edu/dash/sites/osc.hul.harvard.edu.dash/files/Assistance%20Authorization%20Form%20(12-15-11).pdf (accessed September 19, 2012).
44. @mire, http://atmire.com/website/ (accessed September 3, 2012).
45. Richard Holton, "New Open Access Working Group Formed: Formulating Response to Elsevier's Policy Change," *MIT Faculty Newsletter* 24, no. 4 (March/April 2012), http://web.mit.edu/fnl/volume/244/holton.html (accessed September 3, 2012).
46. Armbruster, 7.
47. David Rapp, "Open Access Coalition Formed by 22 Academic Institutions," *Library Journal* (August 8, 2011), www.libraryjournal.com/lj/home/891535-264/open-access_coalition_formed_by_22.html.csp (accessed September 3, 2012).
48. English. Personal communication, February 2012.