

Understanding Faculty to Improve Content Recruitment for Institutional Repositories

From *D-Lib Magazine* Vol. 11 No. 1 (January 2005) <<http://www.dlib.org/dlib/january05/foster/01foster.html>>

Nancy Fried Foster, Lead Anthropologist

Susan Gibbons, Assistant Dean, Public Services and Collection Development
River Campus Libraries, University of Rochester, Rochester, NY 14627

Institutional Repositories and the Adoption Problem

An institutional repository (IR) is an electronic system that captures, preserves, and provides access to the digital work products of a community [1]. In a university setting, an IR may provide a place for faculty work, student theses and dissertations, e-journals, datasets and so on. Whatever the particular focus of the university IR, to be successful it must be filled with scholarly work of enduring value that is searched and cited.

Based on the number of institutional repositories established over the past few years, the IR service appears to be quite attractive and compelling to institutions. IRs provide an institution with a mechanism to showcase its scholarly output, centralize and introduce efficiencies to the stewardship of digital documents of value, and respond proactively to the escalating crisis in scholarly communication [2].

The availability of open-source IR systems has encouraged a proliferation of IRs worldwide, particularly among academic and research institutions. For example, DSpace, developed by MIT Libraries in collaboration with Hewlett-Packard, has been downloaded over 15,500 times since its release in November 2002 [3].

Installing the software, however, is just the first step towards a successful IR. Without content, an IR is just a set of empty shelves. And, in spite of the rapid pace at which organizations are establishing IRs, the quantity of content deposited into them remains quite modest. An April 2004 survey of 45 IRs found the average number of documents to be only 1,250 per repository, with a median of 290 [4]. This is a small number when considering the hundreds of thousands of dollars and staff hours that go into establishing and maintaining an IR. For example, MIT Libraries estimate that their IR will cost \$285,000 annually in staffing, operating expenses, and equipment escrow [5]. With approximately 4,000 items currently in their IR, that is over \$71 spent *per item, per year*.

MIT Libraries certainly have not been complacent in their endeavors to recruit content into their DSpace IR. As part of the DSpace rollout on campus, MIT hired a DSpace User Support Manager in order to work, in part, on content recruitment. When this strategy failed to reap the quantity of documents expected, the skills of a marketing expert were sought [6]. Something seems amiss when even MIT, which has arguably the highest-profiled IR, and which has received national and international press, struggles to recruit content.

The phrase “if you build it, they will come” does not yet apply to IRs. While their benefits seem to be very persuasive to institutions, IRs fail to appear compelling and useful to the authors and owners of the content. And, without the content, IRs will not succeed, because institutions will sustain IRs for only so long without greater evidence of success.

With the generous funding of a 2003 Institute of Museum and Library Services, National Leadership grant, the University of Rochester, River Campus Libraries endeavored to explore the apparent misalignment between the benefits and services of an IR with the actual needs and desires of faculty. Specifically, the grant sought to understand the current work practices of faculty in different disciplines in order to see how an IR might naturally support existing ways of work. This yearlong research has revealed some of the reasons why current IR systems are more useful to faculty in theory than in practice and has resulted in modifications to the University of Rochester's implementation of the DSpace code [7] to better align the repository with the existing work practices of faculty. Moreover, the findings have caused a complete rethinking of how we should explain and promote our IR.

Studying Faculty Work Practices

Our study focused on how faculty members do their research and writing. This entailed looking at how faculty members interact with digital tools and how they organize work in their virtual and physical workspaces. We did this with a work-practice study, a method of fine-grained observation and documentation of people at work based on traditional anthropological participant observation [8].

The best way to conduct work-practice studies is to spend long periods of time with the people under study, observing them as they conduct the usual tasks associated with their work. Work-practice studies typically involve videotaping in order to create a record that can be reviewed repeatedly by the research team [9]. This enables the research team to detect the details that may be missed in the moment of observation and to overcome some of their biases and preconceptions.

In studies of faculty in the university setting, it is not feasible to spend long periods of time observing and taping, so we conducted our observations in brief (usually one hour) sessions. We went into faculty workplaces and videotaped faculty as they did their work, asking them questions and inviting them to show us how they found, used, and disseminated scholarship in digital formats, including both published material and grey literature [10]. We augmented these observations with telephone interviews of faculty and with additional information gathering by librarians. We focused primarily on faculty in the University of Rochester's Departments of Economics, Physics, Political Science, and Linguistics, and the Graduate Program in Visual and Cultural Studies.

Our core team included two librarians, a computer scientist, an anthropologist, a programmer, and a graphic designer. Four reference librarians and a cataloger augmented the core team for project activities, including operating the video camcorder and participating in brainstorming sessions. We believe that including people with different perspectives and expertise made it possible for us to do more research and come up with better ideas, and that the diversity of the team was a key factor in the success of the project.

We conducted and transcribed twenty-five videotaped interview/observations of faculty members. Team members read as many transcripts and viewed as many videos as time allowed. The anthropologist conducted a variety of analyses based on the field data. At some meetings, the whole team analyzed the transcripts together, usually in the context of discussing research questions or performing an activity. For example, we used interview

transcripts to create storyboards of the research cycle from concept to publication for three faculty members, and then returned to the faculty members to discuss and improve their storyboards. We also engaged in spontaneous "blue sky" brainstorming, such as when we imagined that we were faculty members and could magically have and use any tool that would make our research easier or more effective.

We analyzed the data both by rigorous anthropological means and through a more creative, intuitive process. We found that the two methods were complementary and that the combination of the two produced a great number of ideas. Indeed, the result of these sessions was more than 150 different ideas, from general ideas about what faculty members want and need, to specific ideas for possible enhancements to the DSpace software.

We took our large set of enhancement ideas and organized them into categories, for example, "ideas related to version control and other authoring needs," and "ideas related to finding grey literature more easily." The core team evaluated the wide range of ideas, settling on an enhancement to DSpace – specifically, a personal showcase page for each researcher – as the best choice.

Once we had decided upon an enhancement, we began the design work and returned to our faculty participants with prototypes. We videotaped study participants manipulating the prototypes, transcribed and analyzed these sessions, and then incorporated what we learned into our next design phase. Once we had a better design, we made an on-screen prototype using static images that we linked together so that parts of the prototype appeared to work. Conducting usability studies on this "clickable" prototype led to the building of the actual enhancement. This cycle of studying users, designing, and testing prototypes is typical of participatory design and increases the likelihood that the final product will work well for users.

What Faculty Members Want

Our key finding seems obvious in retrospect: what faculty members and university researchers want is to do their research, read and write about it, share it with others, and keep up in their fields. Many of our faculty members are outstanding teachers and some are skilled administrators; they provide service to their departments and fields. But even those who are most committed to the role of professor, broadly defined, complain of overwork, resist clerical responsibility, and resent any additional activity that cuts into their research and writing time.

We found that every one of the people we interviewed used digital tools. Minimally these included email, word-processing programs for authoring, spreadsheets for data storage and manipulation, networks for organizing and storing digital output, online library catalogs and databases for finding and accessing publications, and websites for keeping up and conducting professional activities. However, most interviewees only cared that these tools *did* work; they had little interest in *how* the tools worked or *what* the tools were.

Through analysis of the interviews, we developed a list of individual needs centered mainly on authoring and co-authoring, archiving and disseminating their own work, and finding and reading relevant work by other authors. The people we interviewed want most to be able to...

- Work with co-authors
- Keep track of different versions of the same document
- Work from different computers and locations, both Mac and PC
- Make their own work available to others
- Have easy access to other people’s work
- Keep up in their fields
- Organize their materials according to their own scheme
- Control ownership, security, and access
- Ensure that documents are persistently viewable or usable
- Have someone else take responsibility for servers and digital tools
- Be sure not to violate copyright issues
- Keep everything related to computers easy and flawless
- Reduce chaos or at least not add to it
- Not be any busier

It is essential that anything in an IR be absolutely safe and secure. Beyond that, the single most important criterion of an IR’s value to our faculty members is that *other* people find, use, and cite the work that they put into it. Even the most enthusiastic supporters of IRs will soon lose interest if this criterion is not met.

Once we had identified the needs of faculty members and researchers, we returned to the DSpace features we had cited when initially pitching our DSpace IR to faculty and found that there was no perceived fit between the two.

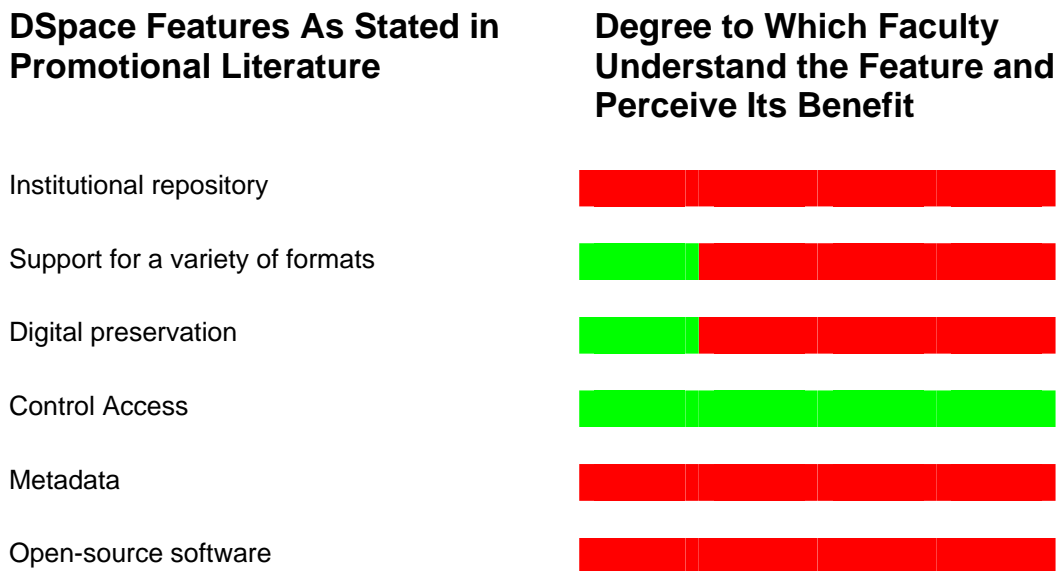


Figure 1. Perceived fit between DSpace features and needs from faculty/researcher perspective. Green indicates understanding while red indicates misunderstanding, lack of understanding, or disinterest.

Faculty members think in terms of reading, researching, writing, and disseminating. They think about the specifics of their research area, whether neutrinos, German film, prosody, or the Congressional Black Caucus. But say “institutional repository” to them, and there is little response.

Most faculty members already assume – usually erroneously – that their materials are rigorously backed up and permanently accessible, even though many have lost data and documents due to crashes or software obsolescence. With a little explanation, they recognize the value of “supporting a variety of formats” and ensuring “digital preservation.” *Every* faculty member understands “control access” and wants it. While some faculty members know what the word “metadata” means in the context of an IR, and all of them can easily understand an explanation of it, the metadata concept is incidental to them and they do not respond to it. “Open-source software” is almost completely meaningless.

Accordingly, when we tried to recruit content using typical IR promotional language, faculty members and researchers did not respond enthusiastically. This is because they did not perceive the relevance of almost any of the IR features as stated in the terms used by librarians, archivists, computer programmers, and others who were setting up and running the IR for the institution. One reason faculty have not rushed to put their work into IRs, therefore, is that they do not recognize its benefits to them in their own terms.

It is beyond the scope of our current project to respond to all of the faculty needs that we identified, particularly those that relate to the authoring of works in progress. DSpace was designed to capture, store, index, and distribute *finished works*, and therefore does not support document versioning or the routing of works among co-authors. Most of the faculty members we interviewed had developed elaborate routines to keep their works in progress safe and organized. For instance, when collaborating with others, they go to great lengths to keep track of the most recent version of a document, developing such makeshift versioning systems as emailing the documents to themselves in order to date stamp each version.

Versioning is also an issue because faculty members often work with documents on more than one computer, across numerous environments. A professor may have a desktop computer in the office, another in the lab, a third at home, and a laptop to carry to the library, café, or conference. If graduate students are assisting in the research, then several more computers are added to the mix. Digital files are being emailed, FTP’ed to departmental and university servers, burned to CD, stored on floppies, zip drives, and USB drives, and even sent to family members out of state for safekeeping.

These faculty members – and we believe them to be typical of faculty at most universities – are in desperate need of an authoring system to assist with document versioning, collaborative authoring, and centralized document access from any computer at any location. This need is so great that an IR for the storage, preservation, and distribution of *finished works* barely registers for them.

Another reason that faculty have expressed little interest in IRs is related to the way the IR is named and organized. The term “institutional repository” implies that the system is designed to support and achieve the needs and goals of the institution, not necessarily those of the individual. Moreover, it suggests that contributions of materials into the repository serve to

highlight the achievements of the institution, rather than those of individual researchers and authors. Our findings strongly support the suggestion of Gandel, Katz, and Metros that the focus should be on the individual, with the emphasis on *personal digital* repositories rather than on *institutional* repositories [11].

When it comes to research, a faculty member's strongest ties are usually with a small circle of colleagues from around the world who share an interest in the same field of research, such as plasma astrophysics or contemporary European critical thought. It is with these colleagues, many of them at other institutions, that researchers most want to communicate and share their work. But most organizations have mapped their IR communities to their academic departments rather than to the subtle, shifting communities of scholars engaged in interrelated research projects. Putting a paper about women in Anglo-Saxon England in the same collection with one on Post-Soviet Kazakhstan just because the authors are in the same academic department may not make sense to faculty because it does not seem that anyone would ever come looking for their work in such a collection. In the absence of a strong connection that would naturally bring these documents together into a collection that other scholars would look for, find, and use, there is no compelling reason for the authors to make the submission.

Enhancing the IR to Meet the Needs of Faculty Users

When we completed our research, we realized that our top priority was to recruit more content for the IR in the short term. While we want to address our faculty's authoring and co-authoring needs, we are deferring work toward those goals until we have increased the amount of content in our IR.

In order to meet our short-term goal, we developed two strategies. One was to try a new strategy for recruiting faculty members, described below, in which we approach them on their own grounds and speak their language. The other was to enhance DSpace to make it much easier for our faculty members to submit their items to the IR and to showcase themselves and their research. Through an iterative, participatory design process, we created two new DSpace elements that enable our faculty members that do just that.

The Researcher Page (Figure 2) is a personalized webpage that we will make available to any University of Rochester faculty member or staff author who puts work into our IR. The Researcher Page will serve as the showcase for all of the researcher's work. Anyone from any computer in the world with an Internet connection should be able to search and find this page and see all the work that a researcher has self-published there. Additionally, the Researcher Page may include links to published work in subject repositories or electronic journals.

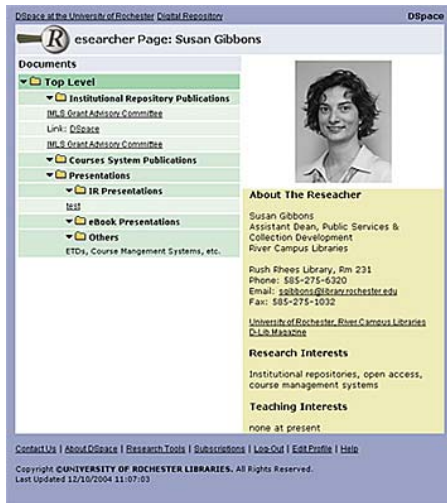


Figure 2. Researcher Page

Supporting the Researcher Page is a Research Tools page (Figure 3). This is the place where the researcher actually completes the tasks of self-publishing and self-archiving. The collocation of the material into collections and the labeling of those collections are completely within the control of the faculty member. In the future, Research Tools will also serve as the homepage for authoring and co-authoring, and it will become the hub for web-based services in support of faculty research. In other words, Research Tools and the Researcher Page will support a wide range of integrated activities that scholars conduct in their various communities.

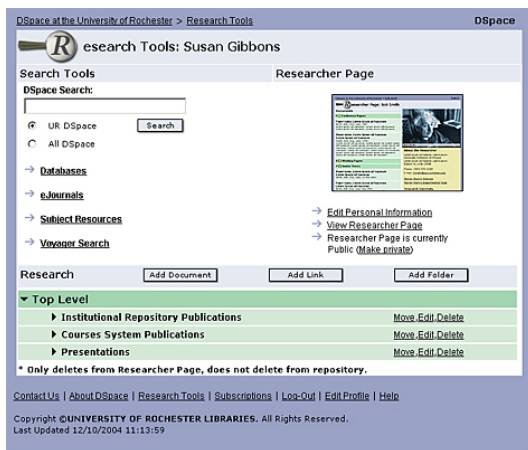


Figure 3. Research Tools

In effect, we are adding a level to the DSpace structure. In addition to the communities and collections that already exist in the DSpace base code, our IR will also have representations for individual community members and their personal collections. The new design adds a strong element of personalization to the system, while retaining the many institutional benefits of an IR.

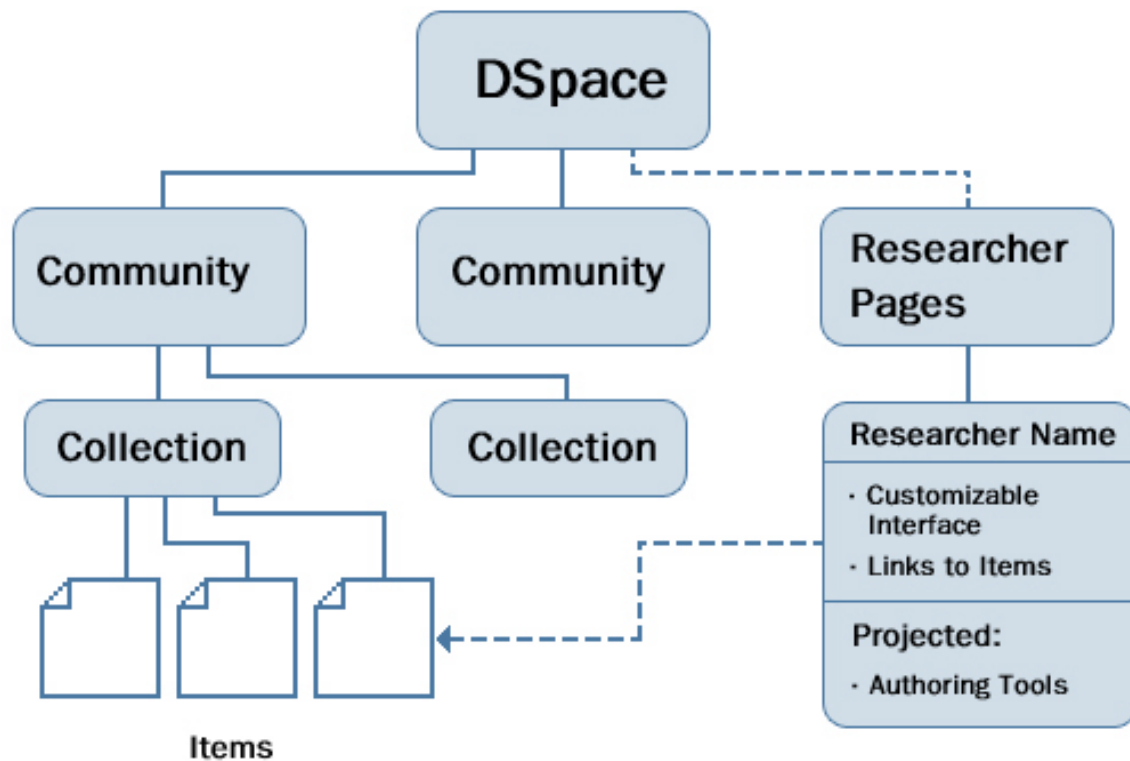


Figure 4. Relationship of Researcher Pages to the Institutional Repository

How We Talk About the IR

With regard to our new approach to faculty, we are now recruiting content for our IR by working with a small early-adopter group and networking from them to their colleagues. We are also creating a wider content-recruitment and user-support structure.

Four users in a small academic department agreed to participate in our research project and evaluate prototypes of our DSpace modifications. They will be the first to have live, Researcher Pages and will participate in further usability testing. As they begin to use these pages, we will monitor their experiences, with a particular interest in the degree to which visitors are searching their pages and viewing their items.

The early-adopter department has close ties to two departments in different fields within our own university and with many departments in the same field at other universities. As early adopters use the new pages with good results, we will explore how they can help us network to other prospective users.

Simultaneously, we are implementing a new content recruitment and user support structure that we hope will make it easier for us to reach out to faculty members, and for faculty members to get the support they need from us, in person and online.

This new structure is based on “library liaisons,” trained subject librarians who are assisting our designated IR collection developer in recruiting content. Library liaisons are available to meet with faculty members individually, or at their departmental meetings to provide information about the benefits of the IR and how it works. Library liaisons and catalogers will also work behind the scenes after faculty members have begun the process of submitting work to the IR by providing support in completing metadata and assigning deposits to appropriate collections.

Rather than approach faculty with a set, one-size-fits-all promotional spiel, these library liaisons operate under the guidance that a personalized, tailored approach works best. As we learned from the work-practice study, what faculty members care most about is their research. Expressing interest in their research, for example by reading a recent article by the faculty member prior to the meeting and then asking a couple of questions about the work, will get their attention and will usually stimulate a very enthusiastic conversation. Throughout the conversation, the library liaison is listening for opportunities to demonstrate how the benefits of the IR respond directly to the faculty member’s web-related research needs. For example, when the faculty member relates frustration over a broken website link, the library liaison can explain that each document in the IR has a unique and stable URL.

By contrast to the language previously used to describe the features and benefits of the IR, we are now describing the IR in language drawn from faculty interviews. Thus, we tell faculty that the IR will enable them to...

- Make their own work easily accessible to others on the web through Google searches and searches within the IR itself
- Preserve digital items far into the future, safe from loss or damage
- Give out links to their work so that they do not have to spend time finding files and sending them out as email attachments
- Maintain ownership of their own work and control who sees it
- **Not** have to maintain a server
- **Not** have to do anything complicated

When we know they are listening, we also tell faculty that if they use selected formats, their files will be persistently usable, or at least viewable, even if they go out of style. We also explain that the IR has a feature that enables them to search for scholarly work in all DSpace-based IRs in Google, a feature of the Researcher Tool page.

In the long run, we envision a system that, first and foremost, supports our faculty members’ efforts to “do their own work” – that is, to organize their resources, do their writing, work with co-authors, and so on. Such a system will include the self-publishing and self-archiving features that the DSpace code already supports, and will rely heavily on preservation, metadata, persistent URLs, and other existing features of DSpace. When we build this system, we will include a simple mechanism for converting works in progress into self-published or self-archived works, that is, moving documents from an in-progress folder into the IR. We believe that if we support the research process as a whole, and if faculty members find that the product meets their needs and fits their way of work, they will use it, and “naturally” put more of their work into the IR.

IRs and Beyond

We are currently building the Researcher Page/Research Tools enhancement in Java and will pilot it with a small group of faculty members. Once we are satisfied with these pages, we will make the code available as open source.

As we roll out the Researcher Page/Research Tools enhancement, we plan to assess whether it is succeeding by focusing on several key outcomes. One is the degree to which users deposit their work into the IR by themselves, using the new pages. Another is whether they begin to give out links to their work in the IR and how well the Researcher Page/Research Tools enhancement works for disseminating their work. We will also be interested in seeing whether we can network from our pilot members to their close colleagues, or whether this networking attempt peters out from lack of interest.

We have begun to apply the work-practice methodology discussed in this paper to other aspects of our work in the library. We are currently conducting a study of undergraduates, focusing on how they do research papers and research-based class assignments. Using a team approach, we will conduct work-practice and related studies to understand how students find and use a variety of academic and non-academic resources, from databases to librarian help to Google. We hope that our findings will help us assist students better through improved reference interviews, better bibliographic instruction, and other improvements indicated by the research.

IRs are at a critical point in their development. While there may be numerous measures for the success of an IR, quantity of content is an obvious and uncomplicated metric. Applying this metric in isolation, it would appear that IRs are failing. However, the findings of our work-practice study suggest that with a faculty-centric approach to the design and marketing of repositories, IRs could become a compelling and useful tool. If properly aligned with the existing practices of faculty, IRs have the potential to fulfill many of their so far unmet expectations.

Nancy Fried Foster
Lead Anthropologist
nfoster@library.rochester.edu

Susan Gibbons
Assistant Dean, Public Services & Collection Development
sgibbons@library.rochester.edu

River Campus Libraries
University of Rochester
Rochester, NY 14627

[1] Clifford Lynch, executive director of the Coalition for Networked Information, describes 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.” “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age.” *ARL Bimonthly Report* 226 (2003).

Available at: <<http://www.arl.org/newsltr/226/ir.html>>

[2] Gibbons, S. (2004). Establishing an Institutional Repository. *Library Technology Report*. 40:4, pp. 11-14.

[3] <http://www.dspace.org>

[4] Ware, M. (2004). Institutional Repositories and Scholarly Publishing. *Learned Publishing*. 17:2, pp. 119.

[5] Barton, M.R. & Walker, J.H. (2002). *MIT Libraries' DSpace Business Plan Project: Final Report to the Andrew W. Mellon Foundation*, p. 33. Available at: <<http://libraries.mit.edu/dspace-mit/mit/mellon.pdf>>

[6] Information shared during presentation by Margaret Branschofsky, “DSpace Marketing and Policy Planning” at the DSpace User Group Meeting, March 10-11, 2004. Available at: <<http://www.dspace.org/conference/presentations/mit-marketing.ppt>>

[7] <https://dspace.lib.rochester.edu/index.jsp>

[8] Two examples of work-practice study are Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity* (New York: Cambridge University Press) and Goodwin, C. (1994). Professional Vision. *American Anthropologist*, New Series, 96:3, pp. 606-633.

[9] Brun-Cottan, F. & Wall, P. (1995). Using Video to Re-present the User. *Communications of the ACM*. 38:5, p. 61.

[10] “Grey literature is that which is produced by government, academies, business, and industries, both in print and electronic formats, but which is not controlled by commercial publishing interests and where publishing is not the primary activity of the organization.” Farace, D. (1997), quoted in Aina, L.O. (2000). Grey Literature and Library and Information Studies: A Global Perspective. *The International Journal on Grey Literature*, 1:4, p. 179.

[11] Gandel, P., Katz, R. & Metros, S. (2004). The ‘Weariness of the Flesh’: Reflections on the Life of the Mind in an Era of Abundance. *Educause Review*, 39:2, pp. 40-51. Available at: <<http://www.educause.edu/apps/er/erm04/erm042.asp>>