SURVEY RESULTS

EXECUTIVE SUMMARY

Introduction

Since 2002, when DSpace and other institutional repository (IR) software began to be available, an increasing number of research libraries and their parent institutions have established institutional repositories to collect and provide access to diverse, locally produced digital materials. This emerging technology holds great promise to transform scholarly communication, but it is still in its infancy.

For the purposes of this survey, an IR was simply defined as a permanent, institution-wide repository of diverse, locally produced digital works (e.g., article preprints and postprints, data sets, electronic theses and dissertations, learning objects, and technical reports) that is available for public use and supports metadata harvesting. If an institution shares an IR with other institutions, it was within the scope of this survey. Not included in this definition were scholars' personal Web sites; academic department, school, or other unit digital archives that are primarily intended to store digital materials created by members of that unit; or disciplinary archives that include digital materials about one or multiple subjects that have been created by authors from many different institutions (e.g., arXiv.org).

In this analysis, the authors have chosen to report the data as percentages based on the relevant number of responses (which can vary by question and within question) unless number totals are clearer. Percent figures have been rounded according to standard rules and they may total to slightly over or under 100%. For questions where respondents supplied textual answers, the authors have only used unambiguous replies in their analysis and calculated percentages accordingly. Respondents who indicated that their institutions currently have IRs are referred to as "implementers." Respondents who indicated they have plans to develop an IR are referred to as "planners."

Background

The survey was distributed to the 123 ARL member libraries in January 2006. Eighty-seven libraries (71%) responded to the survey. Of those, 37 (43%) have an operational IR, 31 (35%) are planning for one by 2007 at the latest, and 19 (22%) have no immediate plans to develop an IR.

One respondent had an operational IR as early as 1999 and a few more came online in 2002 and 2003. Implementation surged in 2004 as 12 repositories became publicly accessible; 14 followed in 2005. Two more were operational in early 2006 and an additional 11 are planned for later this year. Seven others expect their IRs to become accessible in 2007. (One planner indicated that planning and implementation has been ongoing since 2004.)

This data indicate that 30% of all ARL institutions had an operational IR at the beginning of 2006; by the end of 2007 the total may reach at least 55%. While the growth rate appears to be leveling off at this point, IRs will continue to be developed and implemented in the near future.

Implementers and planners are in general agreement about their motivations for starting an IR. The

top three reasons are to increase global visibility of, preserve, and provide free access to the institution's scholarship. These goals are followed closely by a desire to collect and organize the institution's scholarship in a single system (implementers, 89%; planners, 83%). Thirty-eight percent of implementers and 47% of planners were responding to requests for an IR from faculty, staff, and students. Among the other motivating factors was a desire to "Change scholarly communication by demonstrating alternative mechanisms," "Provide a solution to researcher's data management and data publication needs," and "Position the university as a leader in managing digital assets."

All respondents, implementers and planners alike, indicated that the library has been a driving force in the creation of or planning for an IR. Information technology and academic departments advocated for an IR about equally but trailed the library significantly. The administration was an advocate at only about a quarter of the responding institutions.

A wide range of academic units were specifically identified as advocating IRs, such as Aerospace Engineering, Anthropology, Art, Biology, Computer Science, Environmental Studies, Geography, Journalism and Mass Communication, Law, Mathematics, Medicine, Political Science, and Romance Languages. It should be noted that several respondents indicated Graduate Studies or Graduate School Services, with the latter unit explicitly mentioning an interest in electronic theses and dissertations (ETDs).

A variety of other areas on campus were also identified as advocates, such as the Center for Teaching & Learning Excellence, Honors Program, Institute for Policy Studies, Knowledge Media Design Institute, Senate Library Committee, University Archives, and University Press.

Planning, Implementation, and Assessment

Thirty percent of the implementers engaged in planning for six months to a year. Twenty-four percent took from one to six months and an equal number took more than a year to complete the planning stage. The planning process is ongoing for 19%. Only one implementer spent less than a month on the planning stage. More than half of the planners report that this stage is ongoing. For most of the others planning started within the last year.

Roughly a third of the implementers needed less than six months for the implementation phase. Another third took six months to one year to complete the process. Only two needed more than a year. For most of the remaining implementers, that task is ongoing. While almost half of the planners report that they have not reached the implementation phase, the others have either recently entered it or are simultaneously planning and implementing their IRs.

While more than a third of the implementers have not reached the initial assessment phase, yet, for many (43%) assessment is ongoing. A small number (8 or 23%) have completed some assessment. The majority of planners are not ready to assess their IR, but a few report some assessment activity.

Most institutions have conducted or will conduct a pilot project before making their IR public (implementers, 73%; planners, 86%). The pilot project serves multiple purposes. The top two are to determine potential difficulties or problems and plan contingencies and to test processes or procedures. Slightly less important are determining staffing needs (59% and 80%), evaluating and testing IR system options (41% and 80%), estimating costs (41% and 72%), and determining needed material resources (37% and 76%). Other purposes include testing campus interest, building support, and seeding the repository.

Staffing

The overwhelming majority of respondents have appointed or will appoint a project group for planning and implementation (implementers, 92%; planners, 93%). Planning groups range from 2 to

26 members but most are fairly large. The mean number of group members for the implementers is 7.8 and the median is 6; for the planners the mean is 9.1 and the median is 7. Both implementers and planners reported a few groups larger than 15 members.

How library staff are selected to work on an IR implementation varied across the respondents. The top criteria for group members is their functional expertise (94% and 92%), particularly with technology, software, systems, and metadata. Managerial/supervisory expertise is the second criteria for implementers (62%) but only the fifth for planners (35%). Planners report that representation of a particular constituency, such as archives, IT, or faculty, is more important (81%). Being a liaison to a particular stakeholding group, especially faculty, and prior successful project experience are important criteria for a majority of respondents. The least important criteria is workload (15% and 31%), but among the other criteria reported was a willingness to take on additional work.

Survey respondents identified up to four units that were major players in the institution's ongoing IR operations. They indicated the unit name, its responsibilities, the title of the unit manager, the title of the person that the unit reported to, the number of individuals in each staff category (i.e., librarian, other professional, support staff, student assistant, and other), and total FTE in each category.

The data reveal that libraries and their staff are leading the campus IR effort and providing the majority of staffing support for it. There are no campus IT reporting lines for the first units, though there is one joint library-IT reporting line. A small number are found in the second to fourth units and they are often performing server support and similar functions. Out of 58 identified first units, there is only one non-library unit—a graduate studies office.

By far, the most frequent types of units reported for the first unit were digital library/initiatives or systems units within libraries. Also in the list were administrative units, archives, and research or technical services.

It is clear that IRs are a library-wide effort involving many different departments. Aside from technical support units, the lists of second to fourth units includes archives, cataloging, branch libraries, collection development, instruction, metadata, preservation, reference, and special collections. Most of these units report to upper levels of management, such as a library dean or associate/assistant director.

If the mean FTE values for each of the four units are added together, the average number of staff working on an implementers' IR is 28.1. The breakdown by staff category is 7.4 librarians, 7.3 other professional staff, 9.5 support staff, and 3.9 students. The average number of staff working on a planners' IR is 61.2. The breakdown by staff category is 8.8 librarians, 20.8 other professional, 22.2 support staff, and 9.4 students.

Budget

Only 44% of implementers report having a dedicated budget for start-up costs; 48% have a dedicated budget for ongoing operations. Half of the planners anticipate having a dedicated budget for start-up costs and 40% expect to have a dedicated budget for ongoing operations. Many of the respondents who do not have a dedicated budget explained that costs for staff, equipment, etc. were either supported by general library operations already represented within the library's budget, or that existing budget lines were reallocated, or that a consortium or other third party absorbed the costs.

Implementers report a range of start-up costs from $8,000 to $1,800,000, with a mean of $182,550 and a median of $45,000. Planners report a range of $12,000 to $160,000, with a mean of $81,667 and a median of $75,000. The range for ongoing operations budgets for implementers is $8,600 to $500,000, with a mean of $113,543 and median of $41,750. Only two planners knew their budgets for ongoing operations—$100,000 and $133,000, with a mean and median of $116,500.

The distribution of both start-up and ongoing budgets shows concentrations of responses at the

lower and upper ranges, with few in the middle ranges. For start-up, 67% of budgets fall below $75,000, 14% are $75,000 to $125,000, and 19% are $150,000 or greater. The maximum start-up budget ($1,800,000) is far greater than the next highest ($400,000) and is from an institution that included extensive software development and testing costs in its start-up budget.

For ongoing budgets, there is a similar concentration at the ends of the ranges: 50% are below $50,000 and 50% are $100,000 or greater. The maximum ongoing budget ($500,000) is also much greater than the next highest ($300,000) and is reported by an institution that has a major role in a state-wide IR initiative.

The primary method of funding both start-up and ongoing costs is reallocation from existing budgets, but respondents also reported a significant use of new funds. New funds for start-up costs most often came from grants (implementers, 83%; planners, 22%), the parent institution (50% and 33%), one-time supplemental funds (33% and 33%), and other sources (50% and 33%) such as provosts. In almost all cases, reallocated funds are or will be provided by the library (91% and 80%). A few respondents got reallocations from one-time supplemental funds, the information technology department or parent institution, and other sources such as student fees.

New funds for ongoing operations most often are or will be provided by the parent institution, grants, or the library. Reallocating funds is almost always the responsibility of the library (78% and 67%). As with start-up funds, a few respondents got reallocations from one-time supplemental funds, the information technology department or parent institution, and student fees.

Not surprisingly, for the majority of implementers, salaries and benefits account for the largest component of the budget—63% of start-up budgets and 68% of ongoing budgets, on average. This is exceeded only by vendor fees for the small number of institutions whose IR is hosted by an external vendor—70% to 74%, on average. Hardware and software acquisition each account for about a quarter of the start-up budgets and hardware and software maintenance account for under 10%, on average. The allocation for acquisition decreases slightly for ongoing operations and the maintenance allocations increase correspondingly.

Planners allocate the largest percentage of their start-up budgets for hardware acquisition (about 58%, on average) and software acquisition (38%) and a small amount for hardware maintenance (10%). Only one respondent reported a figure for staffing and benefits (57%). For the few planners who have an ongoing operations budget about three-fourths of the budget is allocated for salary and benefits. Much of the rest goes to hardware maintenance.

Hardware and Software

Fifty-three respondents identified the software that is being use to support their IRs. By far, the open source DSpace software is the most common choice of both implementers and planners. Twenty-three of the 33 responding implementers and 14 of the 20 planners (70% each) use DSpace; 20 implementers and 11 planners use it exclusively. Two of the implementers use it in conjunction with CONTENTdm (commercial software); one of these also uses the vendor-hosted DigitalCommons system. One implementer uses DSpace in conjunction with ETD-db and Open Conference Systems (both open source software). Two planners have chosen it in conjunction with open source Fedora software; another with commercial software Digitool.

Of the respondents that don't use DSpace, one implementer uses open source Archimède software and two use commercial CONTENTdm software. Two planners will use open source Fedora software and one will use open source Greenstone software. CONTENTdm, Digitool, and Documentum are the intended commercial systems for one planner each. The ProQuest DigitalCommons system (or the software from the Berkeley Electronic Press

it is based on) is used on all the vendor-hosted platforms.

There is a greater variety of hardware in use. Implementers are about evenly divided between Intel-based servers (Dell models in particular) with either Linux or Windows operating systems and Sun servers with Solaris. Only one uses an Apple Xserve running OS X. All but two of the planners use Intel-based servers, again primarily Dell models, with Linux or Windows operating systems. One uses an IBM RISC server and the other a Sun system.

Memory on the implementers' Sun systems ranges from 2 GB to 64 GB and disk storage ranges from 10 GB to 6 TB. (One institution reported a 12 TB storage unit, but it was not dedicated to the IR.) The Intel-based servers have memory ranging from 512 MB to 4 GB and disk storage ranging from 28 GB to 3 TB. The Macintosh server has 1 TB of storage space. The planners' Intel-based servers have memory ranging from 512 MB to 12 GB and disk storage ranging from 108 GB to 3 TB. The IBM RISC-based server has 4 GB of memory and 275 GB of storage space.

Roughly a third of respondents have made no modifications to the IR software and another third have made minor modifications. About 20% of both implementers and planners have made major modifications to IR software. Implementers are much more likely than planners to have made frequent changes, 22% vs 6%.

Policies and Procedures

Seventy-five percent of implementers and 71% of planners indicated they have or will have written policies and procedures for their IRs. For both groups, 54% have submitted their policies and procedures to an institutional authority for review, or are planning to do so. Most of those who identified the reviewing authority indicated that their policy documents went to the University Counsel.

Respondents place a wide variety of materials in their repositories. Electronic theses and disserta-tions are the most common type of deposit (imple-menters, 67%; planners, 79%). Articles, including preprints and postprints follow closely. The major-ity of respondents include official published ver-sions of articles, conference presentations, technical reports, and working papers. Only a few include university catalogs, yearbooks, or alumni publica-tions.

Only a handful of respondents are actively ne-gotiating with publishers to secure permanent de-posit of e-prints from published serials, but 46% of implementers and 63% of planners are considering doing so in the future.

The widespread inclusion of traditionally un-published material in IRs may reflect the relative ease of recruiting this type of content as well as the fact that these materials in print format do not have robust publishing avenues. Data sets, learning ob-jects, and multimedia materials are the most preva-lent non-traditional materials deposited, with over a third of all respondents indicating they include or will include these materials in the IR. Several respondents mentioned using the IR to house ret-rospectively digitized images and other archival material.

Seventy-four percent of implementers and 83% of planners indicated that they accept any digital file type into the IR, but relatively few (26% and 39%) are committed to functional preservation of every file type. Eighteen percent of implementers and 17% of planners will only accept and preserve specified file types. A few accept certain file types but do not preserve them. Several respondents mentioned following the support levels outlined in MIT's DSpace guidelines (http://www.dspace.org/implement/policy-issues.html#digformats), which include full support and preservation for common file types such as PDF, XML, AIFF for au-dio, and GIF, JPEG, and TIFF for images, among others.

Most deposits to the IR are or will be made by authorized depositors (implementers, 89%; plan-ners, 79%). A significant number of IR staff also

deposit documents for authorized users (78% and 63%). Most respondents are using both methods for deposit rather than one or the other, and many respondents indicated that their deposit procedures are still under construction. Almost all respondents indicated that faculty may deposit their materials in the IR and both implementers and planners allow the work of professional staff, students, and support staff to be deposited, as well. Several respondents also mentioned faculty sponsorship as a means of bringing in work for students or outside contributors.

Fifty-nine percent of implementers and 73% of planners have some method for reviewing documents for copyright compliance or other reasons. In most cases, the review is by individuals outside the IR unit. Only nine respondents report that IR staff review and approve documents. These procedures are not always systematic, but vary by collection and type of document. Additionally, all but six respondents require depositors to sign a deposit agreement.

Sixty-nine percent of implementers and 62% of planners accept multiple versions of the same document. In most cases, versioning is achieved by appending additional files to the original item; most institutions do not permit the depositor to overwrite an earlier version of the document. Eighty-two percent of respondents indicate that IR staff or authors may withdraw documents from the IR. Comments suggest that most institutions permit withdrawal only in cases of copyright infringement or other legal issues. Institutions that permit the withdrawal of documents generally leave a "tombstone"—a reference to the withdrawn document—in the system as a record. Many respondents felt that policy in this area would solidify as they learned more about the legal landscape.

Metadata

Roughly half of the respondents import metadata into their IRs from outside sources, typically by a process of automated mapping from a variety of schemas into Dublin Core. Many are converting data from local schemas and a surprising number mentioned converting records from the MARC format. Ninety-four percent of implementers and 78% of planners allow depositors to enter simple metadata; many of these same respondents also enter metadata on behalf of depositors (implementers, 60%; planners, 56%) or enhance depositor supplied metadata after the fact depending on the material and source (57% and 72%).

Survey results indicate that many institutions are taking significant steps to ensure that their IRs are interoperable with other systems. Ninety-four percent of implementers and 88% of planners indicate that their IR supports the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and a similarly large percentage (89% and 82%) report using persistent identifiers for materials in the IR. This result is consistent with the widespread use of DSpace and other platforms to which these services are integral. Over half of implementers (54%) have ensured that their systems are OpenURL compliant and 41% of planners anticipate having this functionality. Sixty-three percent of implementers have incorporated their IRs into federated search mechanisms. Since only one planner reports doing so, it may be that institutions still in the planning stages have not yet considered how to incorporate the IR with other services, or these institutions may lack federated search tools altogether.

Although most IR platforms in widespread use are OAI-PMH compliant, only one respondent specifically mentioned being crawled by search engines. The issue of optimizing exposure to search engines may become more significant as IRs become more prevalent and stable.

Content Recruitment

The difficulties faced by institutions when recruiting content for their IRs is clearly borne-out by respondents. Only one implementer found recruitment "very easy" and only seven "somewhat easy." Fourteen (40%) found it "somewhat difficult," and

eight (23%) "very difficult;" five (14%) were neutral. In other words, nearly two-thirds of implementers surveyed were sufficiently challenged by the task of recruitment to label their efforts "difficult."

By contrast, about half of the responding planners were neutral. The remainder were evenly divided between "easy" and "difficult."

This difference in perceptions between implementers and planners may reflect both a simple difference in experience and/or the change in perceptions of implementers after an initial recruitment phase of easily identified departmental content. It may be that it becomes increasingly difficult to recruit content after this initial set of objects is added to the IR.

A variety of recruitment strategies are employed or planned by respondents. The majority have tried subject specialist advocacy, identifying likely depositors, presentations to faculty, and offering to deposit electronic materials for authors. Implementers appear to be more aggressive with additional strategies, such as sending electronic announcements, faculty co-recruiting, offering to digitize and deposit printed material for authors, and holding awareness-raising symposia. This practice may indicate that implementers have reacted to recruitment difficulties by trying more and more recruitment strategies.

One recruitment strategy not mentioned above is institutional pressure on authors to submit content to IRs. Only one implementer requires authors to submit content to the IR. One implementer and one planner are considering such a requirement. Half of the implementers and two-thirds of the planners report there is no pressure on authors to submit content. The rest encourage, but do not require, authors to submit content.

Assessment

A small number of implementers (8 or 22%) have conducted research on why users do or do not contribute to the IR; only five planners (28%) have decided to conduct any research. This seems odd since the success of an IR is highly dependant on users contributing to the IR. One explanation for this might be that about a third of the implementers and 71% of the planners answered that they had not yet reached the assessment phase. Because few institutions have conducted assessment of contributor motivation, there is likely to be limited data regarding what factors influence users who contribute to repositories.

While close to 70% of the implementers who have done some form of assessment of the success of the IR have gathered direct feedback from IR users through interviews, surveys, or focus groups, the majority (23 or 79%) have tracked hits on IR content. This is likely due to the fact that it is fairly simple to collect "hit" data from server log files, while the collection and analysis process for more ethnographic user data is significantly more time consuming.

It is clear from the comments that there are many different viewpoints on what constitutes "success" for a repository. One respondent commented about assessing the usability of the interface, while another responded about counting full-text downloads. Clearly, there are many aspects of an IR which need to be examined to determine success.

Current Status of IR

Because the survey respondents have repositories at various stages of development, the numbers of digital objects in the IRs differ significantly. Implementers report a range of 20 objects to over 19,000. Planners report between 4 and 4,500 objects in their repositories. Interestingly, not all the materials stored in the repositories are available to everyone. Forty-four percent of the implementers (16) have material within their repository that is available to only a specific user group, while 36% of the planners (5) intend to restrict access to parts of their IR to specific groups.

Comments from the respondents indicate that there are different reasons for these restrictions as well as different groups to whom use is being re-

stricted. For example, some repositories contain materials that are limited only to campus users, while others have materials that are limited to a specific department or groups of people (such as a specific group of research faculty). Copyright is only one reason that access to materials is limited. Cultural concerns with primary source materials and pending patents were also cited as reasons for restricting access.

Although some institutions restrict access to materials within their repository, few implementers (3 or 9%) supply IR documents to external users for a fee. Primarily, fees appear to be for re-use of images or electronic thesis and dissertations. This seems somewhat incongruous when one considers that 44% of the implementers limit access to materials within their repository. However, there are several possible explanations for this. First, institutions could be storing institutionally licensed materials, such as images, in their IR. Second, the process of collecting per-use fees is missing from several popular open source software packages for IRs. This makes it difficult for institutions to collect fees on a per-use basis without extending the software.

Benefits

Respondents' comments indicate that the top two benefits of IRs are enhanced visibility and increased dissemination of the institution's scholarship (34 responses or 68%) and free, open, timely access to scholarship (23 or 46%). Preservation and stewardship of digital content and preservation of and long-term access to the institution's scholarship are close seconds (18 responses each or 36%), followed by collecting and organizing assets in a central location (12 or 24%). Four respondents (8%) report that another benefit of an IR is the opportunity to educate faculty about copyright, open access, and scholarly communication.

Challenges

Among the top three challenges that respondents face in implementing, promoting, and running an IR are content recruitment/building a critical mass of content (16 responses or 32%), staffing (15 or 30%), and faculty awareness/buy-in/interest/engagement (14 or 28%). Copyright issues and communicating the benefits of the IR to faculty are close behind. Adequate funding and other resources and integrating the staff and workflow of IRs into existing structures were also recognized as challenges.

Conclusion

Based on the survey, what were the major characteristics of operational ARL institutional repositories at the start of 2006? Most IRs had been established in the last two years (or had just been established). By far, the library was likely to have been the most active institutional advocate of the IR. It was also likely to have been the primary unit leading and supporting the IR effort, sometimes in partnership with the institutional information technology unit. The main reasons for establishing an IR were to increase the global visibility of, preserve, provide free access to, and collect and organize the institution's scholarship. In most cases, a project team had been used to plan and implement the IR and a pilot project had been used to determine IR-related issues. If it was not still ongoing, the IR implementation process had most frequently taken six months to a year, with one to six months being the next most common duration.

By a large majority, the most frequently used local IR software was DSpace, with DigitalCommons (or the bepress software it is based on) being the system of choice for vendor-hosted systems. Local systems usually either ran under variants of Linux or Windows on an Intel-based server or under Solaris on a Sun server. A typical IR held about 3,800 digital objects, with ETDs, article preprints and postprints, conference presentations, technical reports, working papers, conference proceedings, and multimedia materials being the most common types of documents. IRs normally support OAI-PMH and, a little over half the time, OpenURL.

Most IRs had written policies and procedures and the majority of them had been submitted to an

institutional authority for review. Faculty members were almost always authorized to directly deposit digital objects in the IR, and professional staff and students were typically able to do so as well. These depositors almost always signed a deposit agreement. Most institutions also authorized IR staff to deposit objects on behalf of users. A deposit review process was common, with documents most frequently being reviewed by department or other institutional officials. Authorized depositors were almost always allowed to enter metadata; IR staff could typically do so as well, plus enhance existing metadata. Most IRs accepted multiple versions of the same document. Document withdrawal was usually possible, but typically had to be done by IR staff under specific circumstances. The vast majority of institutions intended to preserve IR documents, but most of those doing so limited the types of files that would be preserved. Most institutions found IR content recruitment to be somewhat or very difficult and they usually engaged in a variety of recruitment strategies to increase deposits.

The average IR start-up cost had been around $182,500 and its average ongoing operation budget was about $113,500. Reallocated funds from the library's budget were a key source of IR support, as were new funds from grants and the parent institution. Staff had been the largest single IR budget item during start-up and it remained so in ongoing budgets. Many IRs were funded without dedicated budgets, using existing personnel and technical resources.

The typical IR was supported by about 28 FTE from a variety of units within the library and elsewhere, a digital library/initiatives unit managed it, and that unit reported to a high-level library administrator, such as an assistant or associate dean/director. Most institutions modified their IR software to some degree to enhance its functionality.

As one would expect, the perceptions of institutions still planning IRs did not always match the experience of implementers as outlined above, with differences most often occurring over resource and time requirements as well as levels of difficulty. Since these matters can be difficult to accurately project and little data existed at the time the survey was administered that offered guidance, this is not surprising.

Although institutional repositories are at an early stage of development, ARL libraries have demonstrated a strong preliminary commitment to them: 78% of the 87 survey respondents had either implemented an IR or were planning to do so by the end of 2007. Since IRs represent a significant long-term organizational commitment, this is a major expansion of ARL libraries' service role and, along with digital library functions, aptly illustrates how these libraries are rapidly evolving into global digital information providers.

SURVEY QUESTIONS AND RESPONSES

The SPEC survey on Institutional Repositories was designed by the following University of Houston staff members: Charles W. Bailey, Jr., Assistant Dean for Digital Library Planning and Development; Jill Emery, Director, Electronic Resources Program; Anne Mitchell, Metadata Coordinator; Chris Morris, Web Developer 2; Spencer Simons, Director of the O'Quinn Law Library; and Robert Wright, Coordinator of the Pharmacy Library. These results are based on data submitted by 87 of the 123 ARL member libraries (71%) by the deadline of March 9, 2006. The survey's introductory text and questions are reproduced below, followed by the response data and selected comments from the respondents.

Since 2002, when DSpace and other institutional repository (IR) system software began to be available, an increasing number of research libraries and their parent institutions have established institutional repositories to collect and provide access to diverse locally produced digital materials. This emerging technology holds great promise to transform scholarly communication, but it is still in its infancy.

For the purposes of this survey an IR is simply defined as a permanent, institution-wide repository of diverse locally produced digital works (e.g., article preprints and postprints, data sets, electronic theses and dissertations, learning objects, technical reports, etc.) that is available for public use and supports metadata harvesting. If an institution shares an IR with other institutions, it is within the scope of this survey. Not included in this definition are scholars' personal Web sites; academic department, school, or other unit digital archives that are primarily intended to store digital materials created by members of that unit; or disciplinary archives that include digital materials about one or multiple subjects that have been created by authors from many different institutions (e.g., arXiv.org).

This survey is intended to collect baseline data about ARL member institutions' IR activities. The survey authors recognize that there are many possible service models for supporting an IR, ranging from a single, centralized support service provided by a single entity (such as the library) to a much more diffuse support model involving multiple entities (library, campus information technology unit, etc.) playing different cooperative roles. They also recognize that digital archives and repositories are not neat and tidy entities, and they may defy easy categorization. They understand that some institutions have followed carefully prepared IR project plans, some have taken an experimental approach that results in frequent adaptive changes of strategy, and some have done a bit of both. Survey respondents are asked to use their best judgment and try to adhere to the general definition above when responding to the survey. The authors also acknowledge that, in cases where IR responsibility is diffuse, survey respondents may need to gather information from multiple parties. They and ARL thank you for making this extra effort.

BACKGROUND

1. Does your institution have an operational institutional repository (IR) as defined above? N=87

Yes	37	43%
Yes, but institution-wide system is only for article e-prints	0	0%
No, but an IR is in the planning stages	31	35%
No, and we have no immediate plans to develop one	19	22%

If an IR is in the planning stages, please answer the following questions to the best of your ability based on plans at this time.

2. In which year did/will the IR become publicly accessible? N=56

	<2002	2002	2003	2004	2005	2006	2007
N=	1	3	5	13	14	13	7
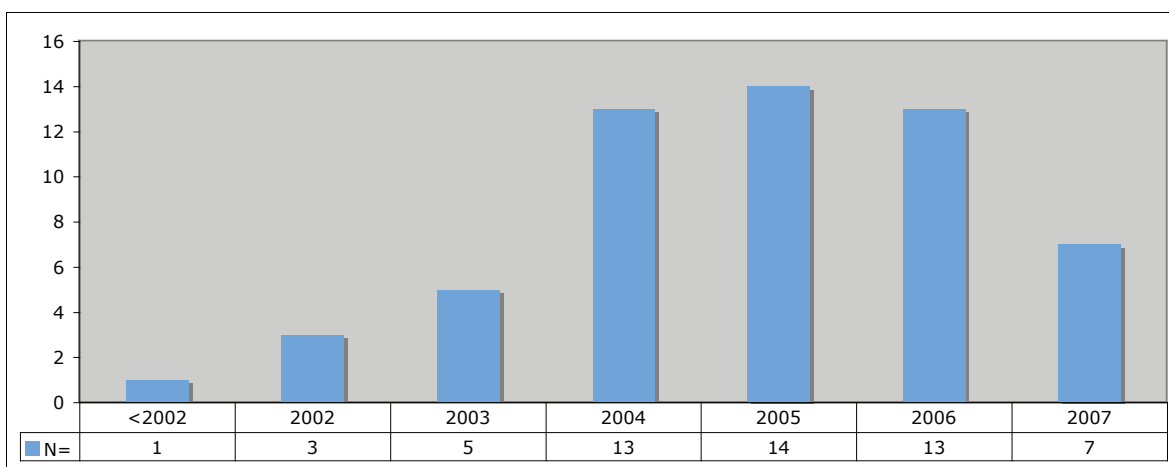
	N	<2002	2002	2003	2004	2005	2006	2007
Have	37	1	3	5	12	14	2	—
Plan	19	—	—	—	1	—	11	7

3. What motivated your institution to establish an IR? Check all that apply. N=67

A desire to:	Total N=67		Have N=37		Planning N=30	
Preserve institution's scholarship	62	93%	35	95%	27	90%
Increase global visibility of institution's scholarship	62	93%	36	97%	26	87%
Provide free access to institution's scholarship	61	91%	33	89%	28	93%
Collect and organize institution's scholarship in a single system	58	87%	33	89%	25	83%
Respond to requests for an IR from faculty, staff, or students	28	42%	14	38%	14	47%
Other motivating factor	15	22%	8	22%	7	23%

Please specify other motivating factor.

Selected Comments from Respondents

Have an IR

"Change scholarly communication by demonstrating alternative mechanisms."

"To make public, materials of value not typically published."

"Provide a source for teaching resources that can be used and/or repurposed."

"Part of developing a distributed institutional repository systems which includes docs and data."

"[Create] centrally accessible repository for educational media and learning objects."

"Increase global availability of [the university's] scholarship."

Planning an IR

"Provide a solution to researchers' data management and data publication needs."

"Foundation for digital library services & infrastructure."

"Apply libraries' leadership to the important issue of preserving and providing access to scholarly communication for the university."

"Manage digital materials for University Archives."

"Contribute to Open Access movement; fulfill legal records retention mandates; position university as a leader in managing digital assets."

"Aggregate library-created digital content. Infrastructure for faculty digitization initiatives."

4. What institutional units have actively advocated establishing an IR? Check all that apply. N=67

	Total N=67		Have N=37		Planning N=30	
Library	67	100%	37	100%	30	100%
Information technology	26	39%	12	32%	14	47%
Academic department(s)	24	36%	12	32%	12	40%
Administration (e.g., president or provost)	17	25%	8	22%	9	30%
Other unit	14	21%	9	24%	5	17%

Academic departments include:

Aerospace Engineering

Anthropology (2 responses)

Applied Economics

Art; Art History; Fine Arts

Biology (2 responses)

Business

Computer Science (4 responses); Computing and Information Science; College of Computing

Civil Engineering (2 responses)

Economics

Education

Electrical Engineering; Electrical and Computer Engineering

Environmental Studies

Geography

Geology

Graduate school services department for electronic theses

Graduate Studies (3 responses)

Industrial Engineering

Information and Library Science

Journalism and Mass Communication

Law

Math

Mechanical Engineering

Medicine

Political Science

Romance Languages

School of Engineering and Applied Science

Veterinary College

Other units include:

Center for Teaching and Learning Excellence

Center for Teaching, Learning, and Technology

Florida Entomologist journal editor

Former Dean of Faculty

Genetic and Public Policy Center

Graduate Studies Office

Honors Program

Institute for Policy Studies

Knowledge Media Design Institute

Learning Technology Centre

Population Center

Research centers/institutes (2 responses)

Resource Centre for Academic Technology

Senate Library Committee

Teaching and Learning with Technology Roundtable

University Archives

University system Office of the President

University Press

PLANNING, IMPLEMENTATION, AND ASSESSMENT

5. In a project of this size and scope, there are typically planning, implementation (e.g., obtaining the IR system, installing it, and customizing it), and initial assessment phases. Please indicate how long it has taken to complete each of these phases in your IR project. N=67

	Planning N=67		Implementation N=62		Initial Assessment N=56	
	Have N=37	Planning N=30	Have N=36	Planning N=26	Have N=35	Planning N=21
Phase not yet reached	—	1	—	12	12	15
Less than one month	1	—	3	2	—	—
One to six months	9	4	8	2	2	1
Six months to a year	11	6	12	1	6	2
More than one year	9	3	2	2	—	1
Ongoing	7	16	11	7	15	2

6. Was there/will there be a pilot project before the IR became/becomes public? N=66

	Total N=66		Have N=37		Planning N=29	
Yes	52	79%	27	73%	25	86%
No	14	21%	10	27%	4	14%

If yes, what was/will be its purpose? Check all that apply. N=52

	Total N=52		Have N=27		Planning N=25	
To determine difficulties or problems that might occur and plan for contingencies	48	92%	23	85%	25	100%
To test processes or procedures	46	89%	23	85%	23	92%
To determine staffing needs	36	69%	16	59%	20	80%
To evaluate and test different IR system options	31	60%	11	41%	20	80%
To estimate costs	29	56%	11	41%	18	72%
To determine material resources needed	29	56%	10	37%	19	76%
Other purpose	16	31%	10	37%	6	24%

Please describe other purpose.

Selected Comments from Respondents

Have an IR
"To test faculty willingness and staff ability to use IR."

"To determine the scope of interest across the institution."

"To build an early-adopter community."

"To test campus interest."

"[To] seed the repository with content before going public."

"[To] gather data for a business plan."

"To craft a list of desired and required features and functionalities in an optimal IR software."

"To test viability of infrastructure to support descriptive and dissemination needs of diversified collection and asset types."

"To build support."

"To identify issues and to determine the feasibility of providing electronic access to theses."

Planning an IR
"[To] determine community willingness to submit materials and identify technical issues."

"To prototype external workflows."

"To build an initial collection of resources to demonstrate value when the IR goes into production."

"To prepare a list of file types we will support and other policies."

"To demonstrate impact of viable IR solution to secure long-term institutional and budgetary support."

"To increase awareness and generate support."

STAFFING

7. Projects of this scale typically require a group (team, committee, task force, etc.) for the initial planning and implementation stages. Did/will your institution appoint a project group to plan and/or implement the IR? N=66

	Total N=66		Have N=37		Planning N=29	
Yes	61	92%	34	92%	27	93%
No	5	8%	3	8%	2	7%

Number of Group Members N=58

	N	Minimum	Maximum	Mean	Median	Std Dev
Total	58	2	26	8.3	6.5	5.6
Have	34	2	26	7.8	6.0	5.7
Planning	24	5	26	9.1	7.0	5.6

Size of IR Project Groups

N=	2	3	4	5	6	7	8	9	10	11	12	13	14	15	>15
	3	4	5	10	7	6	4	1	6	1	2	0	1	2	6
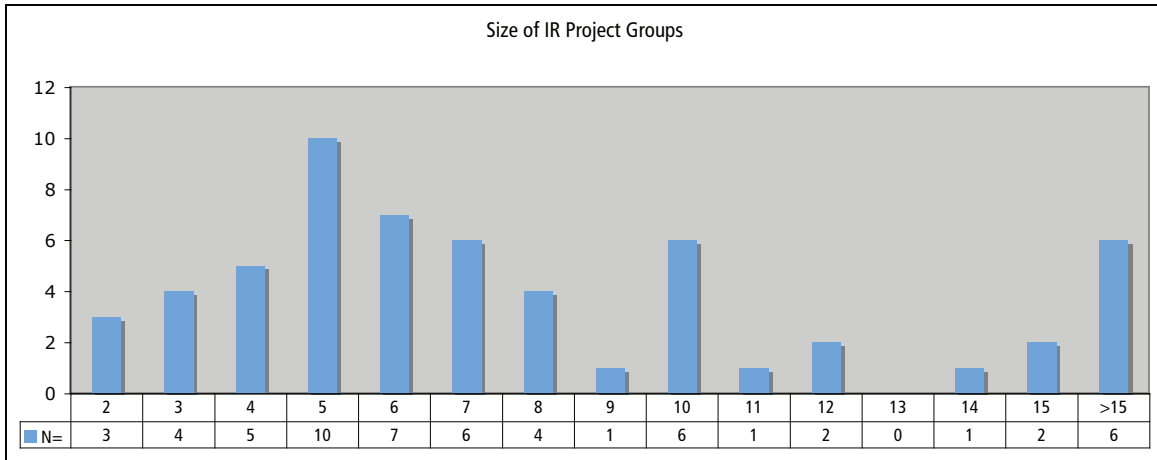
	N	2	3	4	5	6	7	8	9	10	11	12	13	14	15	>15
H	34	3	4	5	3	4	2	2	1	3	1	1	—	1	1	3
P	24	—	—	—	7	3	4	2	—	3	—	1	—	—	1	3

If yes, what criteria were/will be used to select members? Check all that apply. N=60

	Total N=59		Have N=34		Planning N=26	
Functional expertise	56	93%	32	94%	24	92%
Representative of a particular constituency	40	67%	19	56%	21	81%
Liaison to particular stakeholder groups	36	60%	18	53%	18	69%
Prior successful project experience	31	52%	17	50%	14	54%
Managerial/supervisory expertise	30	50%	21	62%	9	35%
Workload considerations	13	22%	5	15%	8	31%
Other criteria	7	12%	4	12%	3	12%

Please describe the criteria.

Functional expertise

Technical expertise (14 responses)

Programming/software development (10)

Systems development and administration (8)

Metadata (6)

Preservation (4)

Collection development (4)

Web technology (3)

Database knowledge (2)

IR software experience (2)

Academic liaison

Archives

Cataloging

Completed initial training with vendor

Copyright

Digital libraries and digital asset management

Digitization

Electronic acquisitions contract experience

e-Publishing software experience

Library technical services experience

Theses submission guideline knowledge

Representative of a particular constituency

Archives (8)

Information technology (8)

Faculty (5)

Library (4)

Collection Development/Subject Specialist (3)

Liaison to faculty/academic departments (2)

Digital projects/library (2)

Administration

Government documents

Graduate Studies Office

Learning technologist

Managers of non-OAI publication database

Senior management

Liaison to particular stakeholder group

Faculty/academic departments (10)

Academic computing

Corporate communications

Dean of Faculty

Graduate students

Library collections

Staff lawyers

Subject specialists

Prior successful project experience

Demonstrated success with other projects, particularly with digital collections, digital library, major computer applications, Web-based information systems, archives

Managerial/supervisory expertise

Able to manage vendor, internal tech staff, and faculty testers

Administrative liaisons in library and central IT

Line and/or task supervision

Management of research

Public services management

Understanding of constituent needs, effective project planning, established/proven record of cross-unit collaboration, communication skills, familiarity with infrastructure, digitization and metadata standards, ability to manage multiple projects/personnel simultaneously

Workload considerations

Hired a full-time programmer/systems administrator

One library assistant is available for IR project

Overlap in membership with related initiatives such as e-research review

Percentage of FTE available for time on task

Pilot project assigned to one of the library's interns

Other criteria

Demonstrated interest in the IR project (4)

Demonstrated interest in scholarly communication issue

Unicode knowledge

Willingness to take on additional work

If no, please explain below how responsibility was/will be delegated for planning and implementing the IR. N=5

Selected Comments from Respondents

Have an IR

"Note, since we adopted the bepress commercial platform (after testing ePrints), most of our work was analysis, spec, testing, rather than architecting and coding. The critical task was managing co-development with vendor based on information gleaned from faculty and CDL testers."

"Core project team; decentralized teams to analyze function issues and to set standards and policies."

Planning an IR

"[Our consortium has] provided, with some participation from their members, a DSpace platform titled 'Research Commons.' So, the platform exists, is public, and is currently providing access. Different schools are in different stages of planning and implementation. We have mounted some important collections without an overall plan, and we now are in critical tipping point to establish credible and useful IR policies."

"Once we agreed on platform, a full time librarian project leader was assigned to fully develop the IR plan and to implement. A technical staff member was also assigned for approximately 30% of her time to assist project leader."

"IR is being deployed by and at our consortium, OhioLINK. Staffing, hardware, and funding are centralized. Steering Committee is state wide. Working committees will be state wide."

8. Please describe the various units (up to four) in your institution that have/will have responsibility for the ongoing operation of the IR. These units may be within the library, the institution's IT unit, or some other unit. It is understood that many units in the library (or elsewhere) will play a part in overall IR support; however, the intent of this question is to identify the major players. Indicate name of the unit, the unit's IR responsibilities, the title of the unit manager, the title of the person to whom the unit reports, the number of individuals in each staff category, and total FTE in each staff category. Please provide any comments that help explain the responsibilities for the ongoing management of the IR. N=58

Unit 1

Have an IR N=34

Name of unit	IR responsibility	Manager's title	Unit reports to
Bibliographic Services	Planning and implementation	Head Bibliographic Services	
Digital Acquisitions Management and Licensing	Digital acquisitions management and licensing	Head	Associate University Librarian for Collection Management and Scholarly Communication
Digital Collections Center—Library Unit	Project management, outreach, repository administrative support	Head, Digital Collections Center	Assistant Director for Technology for General Library System
Digital Development Unit	IR software development and updates	Chief, Research and Digital Development Section	Director, Support and Development Services
Digital Initiatives	Primary IR builder	DI Department Head	Associate Director for Technical Services
Digital Initiatives	Local digital content development (including digitization/imaging)	Coordinator of Digital Content Development	Assistant Dean of Libraries for Scholarly Communication
Digital Initiatives	Marketing, input, education	Head, Digital Initiatives	Dean, University Libraries
Digital Initiatives Group	Special Collections (ContentDM)	Head, Archives & Special Collections	Dean of Libraries
Digital Initiatives Unit	Programming; user needs analysis	Director, Digital Library Initiatives	Three assistant deans
Digital Library and Information Systems Team	Digital libraries, systems, workstations, servers	Team Leader	Dean of the Libraries
Digital Library and Information Technology	Library systems group	Director of Library Systems	AUL for Digital Library and Information Technology
Digital Library Initiative	Maintain, add functionality, work with constituents in using it	Executive Director, Digital Library Initiative	Vice Provost/University Librarian

Digital Library Production	Administer the content component of the project	Head of Digital Library Production	Associate Director for Digital Initiatives
Digital Library Services	Overall planning, management, maintenance,	Head, Digital Initiatives	Associate Dean for Collection and Technology Services
Educational Technology Center	Project coordination; collection management; application administration; archival protocol; metadata and digitization coordination and support	Director	Associate Vice President for University Libraries (chief administrative officer for the University Libraries)
Graduate Studies Office	Establish e-thesis submission regulations, check submitted theses	Director of Graduate Studies Academic Services	Dean of Graduate Studies
Information Technology Services (aka Systems)	Overall co-ordination and implementation; providing technical infrastructure; incubating new approaches	T-Space Service Coordinator	Director, Information Technology Services
IR Working Group	Strategic management and oversight	Project Leader, ScholarlyCommons	Administrative Council (Library's most senior management body)
Libraries Electronic Technologies & Services	Planning, implementation, technology development & support	Manager Digital Library Planning & Development	Associate Director Information Services & Systems
Library	Maintain and enhance DSpace software, identify and add content, publicize	Assistant Head, Digital Library Initiatives	Provost
Library	Everything except backup	Associate Dean for Digital Initiatives	Dean
Library Administration (within)	Management and promotion	Assistant to the Director of Libraries	Director of Libraries
Library Computing and Media Services	Managing all computing for the libraries	Director	Associate Dean
Library Computing Systems	Maintenance and support of Institutional Repository software and hardware	Assistant Director for Library Computing Systems	Assistant Director for Library Computing Systems
Library Information Systems	Management of the software	Assistant to the University Librarian for Information Technology	University Librarian

Library Information Technology	Library information technology	Coordinator, Deep Blue pilot project	Associate University Librarian
Library Systems	Coordination and management of IR	AUL, Library Systems	AUL, Library Systems
Library, Information Technology Services	Support, customization, metadata standards	Associate Chief Librarian, Information Technology	Chief Information Officer
Metadata and Digital Library Services	Running software, updating supporting documentation, meeting with faculty	Head, Metadata and Digital Library Services	AUL for Collections and Access
Scholarly Communication	Management	Coordinator	Dean
Systems Department	Technical support; monitor system trends in ePublishing	eScholarship Program Manager	Head of Systems
Technical Services	Soliciting content, determining viability, creating and reviewing metadata	Assistant University Librarian	University Librarian
University Library Administration	Leadership, direction, strategic planning, promotion	Program Manager	University Library Director

Planning an IR N=24

Name of unit	IR responsibility	Manager's title	Unit reports to
Academic Programs Division	Leadership, planning	Two managers: University Archivist & Forestry Librarian	Associate University Librarian for Academic Programs
Digital Access and Information Architecture Department	Library Web site, OAI, metadata harvesting, etc.	Head, DAIAD	AD for Information Systems and Digital Access
Digital Library Center	Technical, software, some scanning, software specification assistance	Director, Digital Library Center	Associate Director for Technology Services
Digital Library Initiatives	Managing content & especially content intake	Head of Digital Library Initiatives	Deputy Director
Digital Library	Acquire and describe resources; marketing and communication; project management; digital conversion	Director, Digital Library	University Librarian

Digital Library Planning and Development	Project management and technical support		
Digital Library Program	System admin; software development; metadata; infrastructure	Director, Digital Library Program	Jointly, Dean of Libraries and Chief Information Officer
Information Technology	Information technology development and support	Associate University Librarian, Information Technology	University Librarian
Information Technology Services	Installation, configuration, maintenance of D-Space	Interim Manager of Information Technology Resources and Services	Director of Library Services and Information Resources
Libraries/Digital Initiatives	Planning, coordination, and implementation	Digital Initiative Librarian	Associate Dean for Collections
Library		Director Library Technologies & Digital Initiatives	
Library	Plan, implement & support the IR	e-Scholarship Librarian	Head, Digitization Projects with a dotted line to the Director of Libraries
Library Information Technology	Information technology systems development, integration and long term maintenance	Associate University Librarian for IT	University Librarian
Library Information Technology Planning and Policy	Core responsibility for IR	Associate University Librarian for Information Technology Planning and Policy	University Librarian
Library Systems	Working with university IT department to assess technology needs	Associate Dean	Library Dean
Library Systems	Hardware, software, some training.	Head of Systems	AD for Systems and Planning
Library Technology	Technology development and project management	Head, Library Information Systems & Technology	Associate Dean, University Libraries
Library Technology Services	Technical infrastructure		
e-Library	Management and coordination		Acting Director, Information Access & Delivery

Research Services	Liaison with researchers, recruitment of content, articulate services desired by constituency	Head, Research Services	Associate Director for Library Services and Collections
Resources & Services (i.e., the library)	Planning, develop requirements, development, deployment		Senior Associate Dean and Executive Director
Western Archives	Overall coordination	University Archivist	University Librarian

Unit 1

Librarian N=41

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	41	1.0	14.0	2.6	2.0	2.5
Have	26	1.0	7.0	2.6	2.0	1.8
Planning	15	1.0	14.0	2.5	1.0	3.4

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	39	.10	13.5	2.0	1.0	2.4
Have	24	.10	7.0	1.9	1.8	1.7
Planning	15	.10	13.5	2.1	1.0	3.3

Other Professional N=29

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	29	1.0	17.0	2.9	2.0	3.7
Have	19	1.0	17.0	3.1	2.0	4.1
Planning	10	1.0	10.0	2.6	1.5	2.9

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	29	.05	17.0	2.2	1.00	3.9
Have	19	.05	17.0	2.4	1.00	4.3
Planning	10	.10	9.5	2.0	.75	3.0

Support Staff N=24

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	24	1.0	22.0	3.3	1.5	4.6
Have	15	1.0	22.0	3.7	1.0	5.7
Planning	9	1.0	5.0	2.4	2.0	1.7

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	24	.10	19.0	2.7	1.0	4.3
Have	15	.10	19.0	3.0	1.0	5.2
Planning	9	.15	5.0	2.3	2.0	1.9

Student Assistant N=16

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	16	1.0	50.0	7.3	5.0	12.3
Have	10	1.0	12.0	3.9	3.0	3.4
Planning	6	1.0	50.0	12.3	5.0	18.7

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	16	.05	25.0	3.2	1.1	6.0
Have	10	.05	4.0	1.2	.9	1.3
Planning	6	1.00	25.0	6.5	3.1	9.2

Other Staff Category N=6

Have an IR

Graduate Research Assistant (.5 FTE)

1 part-time librarian

Programmer/analyst, UNIX system administrator

Software developer

Volunteers (.5 FTE)

Planning an IR

1 Web site manager, 1 programmer/system administrator

Selected Comments from Respondents

Have an IR

"Assistant Director for Library Computing Systems was co-project director for the Institutional Repository and participated in faculty contact and promotion. Systems staff supported hardware and software installation and maintenance."

"Digital initiatives staff work closely with subject specialists (bibliographers) in promoting the IR to departments, centers, etc."

"The eScholarship Program Manager has dual reporting: reports to Head of Systems and Head of Reference & Instructional Services."

"Head, Digital Initiatives spends about .25 on IR; Application Systems Analyst is being trained to take on some IR duties; student spends approximately 2 hours a week on the IR."

"One programmer working almost 50% of time to develop and support IR and one manager working 10% of time in planning the IR."

"Plan to involve all Cataloging staff by individual subject expertise as the content grows."

"Project management and support for the IR constitutes one aspect of staff effort, however, support for individual collections within the IR fluctuates and involve both principal IR support staff and a variable number of student assistants depending on need. The numbers provided above are an average of ongoing administrative IR effort and project level effort."

"Support staff in the Graduate Studies Office are increasingly devoting time to checking electronically submitted theses rather than print."

"The Digital Projects Librarian has technical responsibility for the project."

"Working Group staff are drawn from public services, collection development, special collections, technical services, IT systems, and administration. Three temporary librarians (Other Professional) provide ancillary project support."

Planning an IR

"One librarian is project leader full time. One librarian in library systems office is backup support to the programmer supporting the project leader. One programmer manages the Digitool software from ExLibris."

"The Library IT unit is the current center of emerging technology integration. We are the main liaison with our consortium, and therefore have the job of integrating the technologies and services made available by them (and others) into first the library, and then hopefully the rest of campus. LIT also is the main contact point for emerging technology that originates from campus IT operations, and we attempt to integrate the library into wider campus IT initiatives."

"There are two full time staff (one library faculty member and one research programmer) dedicated to the library IR. In addition there is one library faculty member (subject specialist) who has 25% of her time designated as a faculty liaison for the IR."

"These numbers are for the whole dept. Only one fulltime support staff and one librarian are working on the IR project."

"This is a new position which is in the process of being hired. The person will be supervised by the Associate Dean for Collections and work closely with Special Collections and University Archives."

"This represents the [main library] staff only. Liaisons from other university libraries will also perform these functions for their constituencies."

Unit 2

Have an IR N=28

Name of unit	IR responsibility	Manager's title	Unit reports to
Applications Group	System administration	Director of Applications Group	Associate Dean for Information Management Services
Archives	Gather and load archival content, esp. campus e-publications	Department Head	Associate Director for Technical Services
Campus IT	Backup	Operations Manager	Head, Campus IT
Central Technical Services	Metadata development	Assistant Director of University Libraries for Technical Services	Associate Vice President for University Libraries (chief administrative officer for the University Libraries)
Collections and Scholarly Communications	Manage print and e-resources	Director	Associate Dean
Collections Management	Planning, implementation, ETD information web pages, liaison with FGS, contact with students, general contact with faculty	Coordinator, Collections Management	Associate Director Collections
Computing Operations & Research Services	Technical advice	Director	Dean
Digital Library Program	Running digital library projects	Head	Associate University Librarian for the Electronic Library
Digital Outreach	Communication with contributors	Digital Outreach Librarian	General Reference

Electronic Publishing Center	Assist with submission of ETDs; large scale scanning for IR; independent digitization projects	Electronic Publishing Librarian	Dean of Libraries
Engineering Library	Library support for the School of Engineering and Applied Science	Head, Engineering Library	AUL for Public Services
Enterprise Academic Systems	Support the Dspace software environment and customize that environment to meet the needs of the scholarly community. We also provide budgetary and administrative support for the ongoing operation of both software and hardware.	Director of Enterprise Academic Systems	Association Vice-Provost for Information Technology
Information Systems and Technology	Maintain server, contribute technical expertise, teach course on electronic submission	Faculty Consulting and Liaison, Client Services	Associate Provost, Information Systems and Technology
Information Technology Services	Provide/maintain server/ storage platform, modify and maintain software	Associate Dean of Libraries	Associate Dean of Libraries
Library	Community coordinator: outreach to faculty at the campus; assistance w/ procedures, mounting content and negotiating copyright clearances	Liaison Librarian	Chief Librarian
Library Computing	Technical support and implementation	Manager, Application and Bibliographic Systems	Director, Library Computing Services
Library Information Technology Support	Technology support	Director of the Digital Library	University Librarian
Library Instruction Research Applications Group— Central IT Unit	Development work, server maintenance, back-up, technical consulting	Head, Library Instruction Research Applications Group	Director of Information Technology
Library Public Services	Faculty contact, public services, policies, promotion, scanning	Assistant Director for Library Public Services	Director of Libraries

Metadata Unit	Metadata support	Head, Metadata Unit	Head, Cataloging and Metadata Services
Reference & Instructional Services	Subject librarians are liaisons with faculty	eScholarship Program Manager	Head of Reference & Instructional Services
Research	Data repository (DSpace etc.)	Associate Dean for Research	Dean of Libraries
Scholarly Resources Integration Department (SRI)	Metadata issues & training	Head, SRI Department	Assistant Director for Technical Services & Information Technology
Server Administration	Computer and administrative support	Chief, Computer and Administrative Support	Director, Support and Development Services
System Integration	Manage the systems component of the application	Chief System Engineer	Associate Director for Digital Initiatives
Systems Office	Systems maintenance	Director of Systems Office	Director of Libraries
Systems, Architecture, and Infrastructure	Keep hardware and operating system running	Director	Vice Provost for Information Technology
University Archives		University Archivist	VP Academic

Planning an IR N=17

Name of unit	IR responsibility	Manager's title	Unit reports to
Collections Access and Management	Electronic theses and dissertations	Director	Associate Dean for Operations and Director of the Library
Collections and Technical Services.	Scholarly communication education	Scholarly Communication Librarian	AD for Collections and Technical Services
Digital Library Center	Data deposit in IR system		
Information Development & Management (IDM)	Planning, develop requirements, development, deployment		Associate Executive Director
Information Technology Division	Platform, implementation	Director of Digital Library Development Laboratory	Associate University Librarian for Information Technology
Library Collection Development	Project leader; policies; content recruitment	Director of Collection Development and Digital Scholarship	Executive Associate Dean, Libraries

Library Digital Programs	Software decisions, ingestion of data, development of technical services	Associate Director for the Libraries Digital Programs	Associate Director for the Libraries Digital Programs
Library Digital Services/Web Administration	Design and organization of IR site; training of IR site users.	n/a	Library Associate Dean
Library ITS	Technical support	Director, Library Information Technology Services	Association University Librarian (Information Services)
Library Special Collections	University archives and library special collections	Head of Special Collections	Associate University Librarian for Collections
Library Systems Department	Support online catalog and other systems	Head, LSD	AD for ISDA
Library Technology Team	Management of technology infrastructure	Chief Technology Officer	Deputy Director
Metadata	IR metadata	Metadata coordinator	
Network Computing Centre	Manage server, handle backups, etc.		Deputy Provost, Chief Information Officer
Services	IR promotion, education, faculty liaison	Associate University Librarian, Services	University Librarian
Systems	Software specification assistance, support Greenstone software	Interim Chair, Systems Department	Interim Assistant Director for Support Services
Systems	Systems	AUL for Systems	University Librarian

Unit 2

Librarian N=30

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	30	1.0	16.0	2.9	1.0	4.1
Have	18	1.0	15.0	3.1	1.0	4.1
Planning	12	1.0	16.0	2.7	1.5	4.2

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	26	.02	16.0	1.9	1.0	3.8
Have	15	.02	13.0	1.4	.5	3.3
Planning	11	.10	16.0	2.5	1.0	4.5

Other Professional N=22

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	22	1.0	12.0	3.0	1.5	2.9
Have	13	1.0	5.0	2.1	1.0	1.6
Planning	9	1.0	12.0	4.3	3.0	3.9

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	20	.05	12.0	2.3	1.0	3.3
Have	12	.05	4.5	.8	.3	1.2
Planning	8	.20	12.0	4.5	4.0	4.2

Support Staff N=16

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	16	1.0	47.0	4.9	1.0	11.4
Have	8	1.0	3.0	1.4	1.0	.8
Planning	8	1.0	47.0	8.5	3.0	15.7

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	14	.05	47.0	5.0	1.0	12.3
Have	6	.08	3.0	.9	.5	1.1
Planning	8	.10	47.0	8.1	2.0	16.0

Student Assistant N=13

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	13	1.0	4.0	2.1	1.0	1.0
Have	7	1.0	4.0	2.0	1.0	1.3
Planning	6	1.0	4.0	2.2	2.0	1.3

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	11	.02	2.0	.7	.50	.6
Have	5	.02	.8	.4	.25	.3
Planning	6	.25	2.0	.9	.75	.7

Other Staff Category N=1

Have an IR
Advisory .01 FTE

Selected Comments from Respondents

Have an IR
"Assistant Director for Library Public Services was co-project director and participated in faculty contact and promotion and developed policies and procedures. One reference librarian reporting to the Head of the Reference Department became DSpace Coordinator, learned how to use the software for input, and worked one-on-one with faculty and department staff. Reference librarians who were subject liaisons to the departments involved also met with interested faculty at the initial meeting, and assisted with metadata. The Head of Digital User Services Department, reporting to the Assistant Director for Library Public Services, and her staff managed staff/students and students scanning reports for departments and research centers. Scanning by the library for the IR later became a part of the project, separately funded by the university."

"Cares for the servers on which the software is loaded."

"Charged to investigate a distributed institutional repository which included both document and data repositories."

"Librarians from both central and special library technical units contribute to the metadata development. The degree of participation varies depending on the number of collections in development. Three primary contributors are represented in these stats plus the director who manages the workflow and has ultimate purview over the metadata construction. An additional cohort of technical services staff serve in an advisory capacity as the Metadata Team. Members of the advisory are called upon to work on projects as needed/appropriate."

"Librarians in other subject areas are beginning to work on the project as the service community expands beyond Engineering. We plan a distributed model for growing the repository. Subject librarians are responsible for identifying 'ripe' opportunities for new IR content, advocating the IR to campus constituencies, and developing linkages between the IR and content holders."

"Only two people are involved in the IR, one librarian and one support staff each working 10% of their time. The librarian involved is also the Coordinator of Collections Management. The support staff member does most of the ongoing work with students and Web pages."

"The computer analyst involvement in this project is the security and maintenance of the server."

"This includes the various library liaisons to faculty department."

"This is system administration and the technical aspects of managing the application itself (installation, upgrades, patches, etc.)"

Planning an IR

"One of the programmer analysts is in library school."

"Special Collections has been a very good leader in the original usage of the IR—they have done the most work to date in the establishment of interim meta-data standards and policies for collection inclusion. The departure of the head of special collections in fall 2005 has stalled our efforts at effective policy creation and implementation. In concert with LIT, Special Collections is critical to the useful integration of any IR activities. However, whether they 'own' the IR is very much up in the air."

"We expect that only the librarians will be active in promoting and teaching the IR."

Unit 3

Have an IR N=15

Name of unit	IR responsibility	Manager's title	Unit reports to
Administration	Environmental scanning; university administration liaison	Associate Dean for Public Services & Collection Development	Dean of Libraries
Archives & Special Collections	Planning	Head, Archives and Special Collections	Director of Libraries
Archives and Special Collections	Input content	University Archivist	
Collection Services	Collection/Subject librarians are liaisons with faculty	Head of Collection Development	Associate University Librarian, Collection Services
Library Technical Services	Provide metadata for selected collections	AUL	Director of Library
Information Technology Department	Document repository (Digital Commons)	Assistant Dean for IT	Dean of Libraries
Library	Coordinate project, maintain upload submission forms, metadata database and public interface	Liaison Librarian	University Librarian
Library Systems	Server administration & support	Supervising Programmer/ Analyst	Associate Vice President for University Libraries (chief administrative officer for the University Libraries)
Office of Information Technology	Server administration, batch data loading, troubleshooting	Senior Systems Manager	Chief Information Officer

Public Services	Outreach and marketing, some user support	Associate Director for Public Services	Director of Library
Research Services	It will be to work with faculty to acquire content	Assistant Head of Research Services	Associate Director for Research Services
Resource Acquisition and Description	Metadata	Head	Associate Dean of Libraries
Systems	Manage library support software and servers	Head of Systems	Associate Director for Technical Services
University IT Services	IT	Director of IT Services	VP Services
University Library	Mounting content and negotiating copyright clearances	Systems Librarian	Chief Librarian

Planning an IR N=10

Name of unit	IR responsibility	Manager's title	Unit reports to
Collection Resources	Providing metadata	Associate University Librarian, Collection Resources	University Librarian
Information Technology Services	Maintaining server space; coordinating with Library Systems		University CIO
Library Administration	Liaison to faculty/university administration	University Librarian	Provost
Library Web Services	IR & Web interface support as needed	Head, Library Web Services	Assistant Dean for Systems
Preservation	Coordinate implementation group, write documentation, establish policies and procedures	Chair, Preservation Department	Associate Director for Technology Services
Public Services.	Content recruitment, training.	Interim AD for Public Services.	Interim AD for Public Services.
Reference Department	User support; training for other library staff	Head, Reference Department	Director of Public Services, Libraries
Scholarly Communication	Copyright, permissions, education and developing new issues related to institutional repositories and scholarly communication	Endowed Chair	Dean

Technical Services	Metadata		
Technical Services	Development & application of metadata	Head, Technical Services	Associate Director for Library Services and Collections

Unit 3

Librarian N=21

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	21	1.0	23.0	4.1	1.0	5.6
Have	12	1.0	23.0	4.3	1.5	6.5
Planning	9	1.0	14.0	3.8	1.0	4.5

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	19	.03	23.0	3.6	1.0	5.8
Have	11	.01	23.0	3.7	.5	7.2
Planning	8	.50	10.0	3.5	1.0	3.5

Other Professional N=7

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	7	1.0	39.0	7.9	2.0	13.9
Have	4	1.0	7.0	3.0	2.0	2.8
Planning	3	2.0	39.0	14.3	2.0	21.4

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	7	.1	39.0	7.3	2.0	14.2
Have	4	.1	7.0	2.0	.4	3.4
Planning	3	2.0	39.0	14.3	2.0	21.4

Support Staff N=7

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	7	1.0	25.0	7.9	2.0	10.4
Have	5	1.0	25.0	6.2	2.0	10.5
Planning	2	3.0	21.0	12.0	12.0	12.7

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	7	.05	24.5	7.4	2.0	10.5
Have	5	.05	24.5	5.6	1.0	10.6
Planning	2	3.0	20.5	11.8	11.8	12.8

Student Assistant N=6

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	6	2.0	14.0	6.7	6.5	4.6
Have	4	2.0	9.0	5.3	5.0	3.8
Planning	2	5.0	14.0	9.5	9.5	6.4

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	6	.08	4.0	2.2	2.5	1.5
Have	4	.08	4.0	2.3	2.5	1.8
Planning	2	1.00	3.0	2.0	2.0	1.4

Other Staff Category N=0

Selected Comments from Respondents

Have an IR

"Five percent of one librarian is dedicated to the repository."

"Head of Archives involved in IR planning but limited involvement in ETD and LOR pilot projects."

"Large department with many staff. They are providing supervision for the IR's programmer analyst, project staffing for development work, networking, and security and maintains of server, etc."

"Regular environmental scans of IR use and adoption in higher education; promotes use of IR at university administration level."

"The total FTE is not static. During times of software adjustment the percentage of FTE can be much higher. The 5% reflects the amount of time devoted on average in normal times."

"We are in the planning stages of implementing this aspect of the program. The Total FTE figure above is just an estimate at this time."

Planning an IR

"Expect only some staff will be involved in IR."

"We have only one librarian actually working on metadata."

Unit 4

Have an IR N=9

Name of unit	IR responsibility	Manager's title	Unit reports to
Bibliographic Control	Planning, metadata + controlled vocabularies	Head, Bibliographic Control	Associate Director Collections
Cataloging	Metadata creation and editing	Metadata cataloger	Head of Cataloging
Cataloging and Metadata Services	Metadata management	Head of Cataloging and Metadata Services	
Fine Art Library	Mounting content and negotiating copyright clearances	Fine Art Librarian	Department Head, Department of Fine Art, Faculty of Arts and Science
Information Technology Division	DSpace administration, account management, testing, troubleshooting, news releases, providing support for IR communities	Assistant Director for Technical Services & IT	Assistant Director for Technical Services & IT
Archives	Promote use of DSpace for archival collections	Head, Archives	Associate Director for Collections
Reference	Faculty liaison	Head of Reference	Associate Dean for Public Services & Collection Development
Special Collections	Selection of collections for addition to the repository; development of collections (digitization & metadata creation)	Director of Music Library & Special Collections	Associate Vice President for University Libraries (chief administrative officer for the University Libraries)
Technical Services	Metadata; digital collections end-processing	Head of Cataloging; Head of Acquisitions	Associate University Librarian, Collection Services

Planning an IR N=5

Name of unit	IR responsibility	Manager's title	Unit reports to
Cataloging Department	Metadata; mediated submissions; e-dissertations	Head of Cataloging Department	Director, Technical Services, Libraries
Electronic Resources	IR licensing issues and integration into electronic resources program		

Office of Libraries Technology	Hardware, software and system support	Director, Office of Libraries Technology	Dean
Special Collections/Archives	Liaison to institutional depositors		
University Archives	Liaisons with faculty and university	University Archivist	Head, Special Collections and Area Studies Department

Unit 4

Librarian N=14

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	14	1.0	5.0	1.6	1.0	1.1
Have	9	1.0	5.0	1.9	2.0	1.3
Planning	5	1.0	2.0	1.2	1.0	.5

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	12	.04	2.00	1.1	.25	.67
Have	8	.04	1.25	.4	.25	.42
Planning	4	.05	2.00	.7	.30	.91

Other Professional N=3

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	3	1.0	3.0	2.0	2.0	1.0
Have	2	2.0	3.0	2.5	2.5	.7
Planning	1	1.0	—	—	—	—

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	3	.5	3.0	1.5	1.1	1.3
Have	2	1.1	3.0	2.1	2.1	1.3
Planning	1	.5				

Support Staff N=1

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	1	3	—	—	—	—
Have	1	3	—	—	—	—
Planning	—	—	—	—	—	—

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	—	—	—	—	—	—
Have	—	—	—	—	—	—
Planning	—	—	—	—	—	—

Student Assistant N=3

# of Staff	N	Minimum	Maximum	Mean	Median	Std Dev
Total	2	1.0	1.0	1.0	1.0	—
Have	1	1.0	—	—	—	—
Planning	1	1.0	—	—	—	—

Total FTE	N	Minimum	Maximum	Mean	Median	Std Dev
Total	3	.07	6.0	2.1	.25	3.4
Have	2	.07	6.0	3.0	3.00	4.2
Planning	1	.25	—	—	—	—

Other Staff Category N=0

Selected Comments from Respondents

Have an IR

"One librarian spends 1/3 of her time training and content recruitment; four other librarians incorporate IR content recruitment as part of their collection development responsibilities."

"The Electronic Resource Cataloguer is involved in IR planning and metadata standards for 5% of her time."

"We are still in the planning stages with respect to how best to employ staff from this area."

9. Is there/will there be a dedicated budget for start-up costs and/or the ongoing operation of the IR? N=58

Start-up Costs N=56

	Total N=56		Have N=34		Planning N=22	
Yes	26	46%	15	44%	11	50%
No	30	54%	19	56%	11	50%

Ongoing Operation N=48

	Total N=48		Have N=33		Planning N=15	
Yes	22	46%	16	48%	6	40%
No	26	54%	17	52%	9	60%

Start-up Costs N=21

	N	Minimum	Maximum	Mean	Median	Std Dev
Total	21	$8,000	$1,800,000	$153,726	$48,000	387,217
Have	15	$8,000	$1,800,000	$182,550	$45,000	458
Planning	6	$12,000	$160,000	$81,667	$75,000	54,647

	Total	Have	Planning
<$25,000	5	4	1
$25,000–49,999	6	5	1
$50,000–74,999	3	2	1
$75,000–99,999	1	1	—
$100,000–124,999	2	—	2
$125,000–149,999	—	—	—
≥$150,000	4	3	1

Ongoing Operation N=16

	N	Minimum	Maximum	Mean	Median	Std Dev
Total	16	$8,600	$500,000	$113,913	$71,750	133,494
Have	14	$8,600	$500,000	$113,543	$41,750	143
Planning	2	$100,000	$133,000	$116,500	$116,500	23,335

	Total	Have	Planning
<$25,000	4	4	—
$25,000–49,999	4	4	—
$50,000–74,999	—	—	—
$75,000–99,999	—	—	—
$100,000–124,999	3	2	1
$125,000–149,999	2	1	1
≥$150,000	3	3	—

Selected Comments from Respondents

Have an IR

"During the ETD pilot project, a programmer was needed full time and the Faculty of Graduate Studies and the libraries split the start-up costs for the programmer salary. The IR is still under development and the programmer is currently spending almost 50% of his time on IR development and support."

"The amount indicated for start-up cost is for hardware; we do not have a figure for salary cost during that period. The amount indicated for ongoing operation is for salaries only."

"Costs are estimates (some ongoing costs not known at this initial point of startup.)"

"Does not include existing staff costs for IR coordination, marketing & liaison with faculty, training. These activities will be absorbed into existing library responsibilities."

"Ongoing operation will, for now, be integrated in the regular operation budget of the library."

"Operations to be streamlined into normal workflow procedures."

"Funds for a server, one year of entry level librarian's salary (reference librarian DSpace Coordinator) student wages, and publicity for a total start-up cost of $150,000."

"Includes 24/7 technical support, site set up for faculty units, training of unit administrators, report generation, etc."

"Costs are for the license for system plus high level of staff currently managing the system. Staff costs could be less. Some academic departments use students to help faculty deposit materials."

"Start-up IR costs have been absorbed within existing University Libraries budgets. Ongoing supports are budgeted for application licenses. Personnel and costs for the expansion of the server array are under

discussion."

"This project was done without a dedicated budget."

"The IR, which consists of both DSpace and CONTENTdm, does not have a separate budget, but is included within the digital library and IT budgets."

"Budget concerns are assumed under overall library budget by the director."

"IR is part of library's repository development program and was not separately budgeted."

"Funded via library's operational budget."

"No dedicated budget at this point. We typically don't do programmatic budgeting for operations. We fund server/storage hardware out of ITS ops budget, so any incremental increases to server/storage needed to support IR are handled that way."

"One limited term appointment to provide leadership and coordinate projects which create digital collections through digitization of electronic publishing."

"The School of Engineering and Applied Science funded the pilot, the library provided staff. Cost of the service has since been assumed by the library to keep budget from inhibiting adoption by other schools and centers. We expect academic units to contribute a larger portion of the staffing need as the IR grows."

Planning an IR

"Too early to indicate the costs."

"Since IR metadata and digital object content will (probably) reside within Greenstone, many funds dedicated towards creation of local digital library are also in support of IR. While the costs in terms of programming and development staff have been significant, to pull out $ for IR alone would be difficult and time consuming."

"As the IR we currently have available is hosted and managed by the consortium, the original start-up costs were absorbed by the yearly regular membership fees of the consortium members. Ongoing operation has not yet been effectively budgeted, but as we move more and more into predicted heavy usage, a cost model needs to be established."

"The ongoing operation budget is approximate only. This is a six-year project funded through the Office of the Provost, the University Library, and the Academic Computing department. Part of the project is to determine the true costs of operating an institutional repository for the university."

"The decision about the exact amount that will be allotted for start-up costs and ongoing operation has not been determined by the Library Dean and the University CIO."

"Out of pocket expense has been limited to desktop development tools and server/storage/tech support costs."

"A Provost Opportunity Fund grant was awarded for start-up costs. Ongoing operations are a part of the library's budget."

"Budget is centralized at our consortium."

"Amount for new server, SAN storage array purchase."

"Work will be incorporated into existing responsibilities; some evaluation work funded with grants; one new staff person."

"We don't have the business model worked out at this stage. We do have some money ($100,000) recurring dedicated through an odd channel that is paying for some technology and liaison staff. Our initial hardware is being paid for through $50,000 of equally opportunistic oddball funding."

"Start up costs have not yet been determined; a dedicated budget would be considered only after the pilot project is completed and cost estimates are available."

"We will be looking for grant funding."

If yes, what is the source of the funds for the budget? Check all that apply. N=33

Start-up Costs N=29	New Funds N=15			Reallocation N=21		
	Total N=15	Have N=6	Planning N=9	Total N=21	Have N=11	Planning N=10
Parent institution	6	3	3	2	—	2
Library	2	1	1	18	10	8
Information Technology	2	1	1	2	—	2
One-time supplemental funds	5	2	3	3	1	2
Grant	7	5	2	—	—	—
Other source	6	3	3	4	2	2

Ongoing Operation N=28	New Funds N=15			Reallocation N=24		
	Total N=15	Have N=9	Planning N=6	Total N=24	Have N=18	Planning N=6
Parent institution	7	5	2	4	2	2
Library	6	3	3	18	14	4
Information Technology	2	1	1	3	2	1
One-time supplemental funds	2	—	2	1	—	1
Grant	4	4	—	—	—	—
Other source	2	2	—	3	3	—

Please describe other source.

Start-up Costs, New Funds

    Atlantic Foundation

    Faculty of Graduate Studies

    Personal gift from provost to the library

    State IT grant to consortium

    IT Grant purchased software

    University Provost

Start-up Costs, Reallocation

    Student fees

    ETC, Library Systems, and CTS budgets

    ILS hardware reimbursement

    Library operational budget

Ongoing Operation, New Funds

    "eScholarship Repository is funded directly from the UCOP."

    University Office of the CIO

Ongoing Operation, Reallocation

    Student fees

    ETC, Library Systems, and CTS budgets

Please estimate the percentage of the budget allocated to each of the following categories. N=26

Start-up Costs

Total N=21	N	Minimum	Maximum	Mean	Median	Std Dev
Staffing and benefits	13	.25	100	62.8	65.0	33.6
Hardware acquisition	14	.75	100	37.3	27.5	34.6
Software acquisition	6	5.00	90	30.7	17.0	33.5
Hardware maintenance	6	1.00	20	9.6	10.0	7.2
Software maintenance	3	2.00	10	5.3	4.0	4.2
Vendor fees (if IR is hosted by an external vendor)	5	2.00	100	70.2	99.0	43.8
Other category	3	1.00	29	—	—	—

Have an IR N=16	N	Minimum	Maximum	Mean	Median	Std Dev
Staffing and benefits	12	.25	100	63.3	67.5	35.2
Hardware acquisition	9	.75	100	25.6	17.0	31.6
Software acquisition	3	5.00	50	23.0	14.0	23.8
Hardware maintenance	3	2.50	15	9.2	10.0	6.3
Software maintenance	2	2.00	10	6.0	6.0	5.7
Vendor fees (if IR is hosted by an external vendor)	5	2.00	100	70.2	99.0	43.8
Other category	2	1.00	29	—	—	—

Planning an IR N=5	N	Minimum	Maximum	Mean	Median	Std Dev
Staffing and benefits	1	57	—	—	—	—
Hardware acquisition	5	10	85	58.4	75.0	31.9
Software acquisition	3	5	90	38.3	20.0	45.4
Hardware maintenance	3	1	20	10.3	10.0	9.5
Software maintenance	1	4	—	—	—	—
Vendor fees (if IR is hosted by an external vendor)	—	—	—	—	—	—
Other category	1	1	—	—	—	—

Please describe other category.

Marketing (1%)

Travel (1%)

Expertise System Development (29%)

Ongoing Operations

Total N=20	N	Minimum	Maximum	Mean	Median	Std Dev
Staffing and benefits	18	10	100	69.2	79.5	29.1
Hardware acquisition	3	15	35	23.3	20.0	10.4
Software acquisition	2	10	19	14.5	14.5	6.4
Hardware maintenance	8	2	25	11.3	11.5	7.4
Software maintenance	3	3	20	12.7	15.0	8.7
Vendor fees (if IR is hosted by an external vendor)	5	34	100	73.8	80.0	27.2
Other category	5	4	55	—	—	—

Have an IR N=18	N	Minimum	Maximum	Mean	Median	Std Dev
Staffing and benefits	16	10	100	68.3	80.5	30.8
Hardware acquisition	3	15	35	23.3	20.0	10.4
Software acquisition	2	10	19	14.5	14.5	6.4
Hardware maintenance	6	2	25	10.3	7.5	8.5
Software maintenance	2	3	20	11.5	11.5	12.0
Vendor fees (if IR is hosted by an external vendor)	5	34	100	73.8	80.0	27.2
Other category	4	5	55	—	—	—

Planning an IR N=2	N	Minimum	Maximum	Mean	Median	Std Dev
Staffing and benefits	2	70	83	76.5	76.5	9.2
Hardware acquisition	—	—	—	—	—	—
Software acquisition	—	—	—	—	—	—
Hardware maintenance	2	13	15	14.0	14.0	1.4
Software maintenance	1	15	—	—	—	—
Vendor fees (if IR is hosted by an external vendor)	—	—	—	—	—	—
Other category	1	4	—	—	—	—

Please describe other category.

Travel (4%)

Production services (15%)

Refresh/preservation (20%)

Server farm charges, storage farm charges, network charges (55%)

If there is not a dedicated budget for the operation of the IR, please describe how operational costs are covered. N=33

Selected Comments from Respondents

Have an IR

"Operating costs are covered by redirecting staff efforts and by allocating a part of our budget to cover operating expenses."

"Hardware and technical support costs are part of existing technology budget that covers development, acquisition, maintenance and support."

"We are using open source technology which is hosted on existing servers. The maintenance is done by our regular staff."

"Portions of various staff members time reallocated to the IR project. Full time programmer salary and benefits came from not filling an empty position within public services."

"Costs for staff and vendor are reallocated from other uses in the libraries."

"Personnel, hardware, and software costs are currently represented within library unit budgets."

"Operational costs are covered by absorbing the work into existing units and activities."

"Annual licensing of system funded by Information Technology Services (ITS); reallocation of library resources for all other IR operational costs."

"IR is operated out of system-wide office, not our institution."

"IR is a department within the library funded in the same way as other departments."

"We use free, open-source software; all other costs (server, staff time) are part of our standard, over-all budget."

"Technical development and support are done by existing staff charged with repository and digital services development and support (librarians and programmers). Collection selection and digitization are done by existing staff (librarians) charged with developing the library's digital collections. Similarly, metadata services are provided by existing cataloging staff."

"Because of the intermingling of software and hardware acquisition and maintenance for other projects, it is hard to identify IR costs. However, staffing is by far the largest cost."

"Absorbed by operations budgets of the three campus departments participating (Graduate Studies, Information, Systems, & Technology, and the library)."

"Software licensing fee is paid from library technology fee; EPC is partially self-supporting through grants and contract work; staff responsibilities in DLS have been shifted to allow librarians in that department to take on this new work."

Planning an IR

"At this time, we are still in the planning stages. We have committed staff time to the Greenstone programming and to the planning group. Dedicated funds were used to purchase hardware for the DL (Digital Library) which will include the IR. There are no hard IR funding figures yet."

"As mentioned above, the current consortium membership fees cover the IR operations. But this is not a long-term model."

"Reallocation of existing librarian position; commitment of systems and reference librarians/staff to the project."

"Operational costs of the pilot project will be covered by existing area budgets."

"At this time expenses for the IR are woven into the operations of the involved units. Three servers are

involved, a portion of one is dedicated for library use, another is part of a federal grant project, and a third is part of a state grant project."

If you offer fee-based IR services (e.g., digitization or metadata enhancement) that provide IR income, please identify them. N=1

"We do offer digitization services (fee-based) to departments. At this time we are not using these funds for the IR."

HARDWARE AND SOFTWARE

10. Please indicate whether your IR system is located at your institution or consortium (e.g., DSpace) or at an IR vendor's site (e.g., DigitalCommons). If it is locally hosted, briefly describe the IR's central site hardware and software configuration, identify the IR software used, the computer model, computer memory, disk storage, and operating system. If it is vendor hosted, please identify the vendor. N=59

	Total N=59		Have N=36		Planning N=23	
Locally hosted by parent institution	49	83%	28	78%	21	91%
Locally hosted by consortium	3	5%	1	3%	2	9%
Hosted by vendor	7	12%	7	19%	—	—

Locally Hosted IR N=48

Have an IR N=28

Freeware	Commercial Software	Computer Model	Memory	Storage	Operating System
Archimède		Compaq Proliant ML 370 Dual P3 750Mhz	1 GB	356 GB	Windows 2000
DSpace (6)					
DSpace				.5 TB	
DSpace				500 GB	Linux
DSpace		Apple XServe		1 TB	OS X
DSpace		Dell PowerEdge 1650	1 GB	36 GB	SUSE Linux 9

DSpace		Dell PowerEdge 2550		36 GB	Red Hat Linux
DSpace		IBM X306	512 MB	2 x 40 GB	Unix
DSpace		ProLiant DL360 G3	2 GB	10 TB	Linux
DSpace		SunFire 280R	4 GB	72 GB	Solaris
DSpace		SunFire 480	4 GB	300 GB	Solaris
DSpace		Sun 240	2 GB	4 x 146 GB	Unix
DSpace		Sun 880 X 2	64 GB	10 GB	Solaris 10
DSpace		Sun 440's	16 GB	6 TB	Solaris
DSpace		Sun 440	16 GB	12 TB	Solaris
DSpace		SunFire 280R	4 GB	~500 GB	Solaris 8
DSpace, ETD-db, Open Conference Systems		Sun 480		438 GB	Solaris
DSpace	CONTENTdm	Compaq DL380	2.5 GB	280 GB	SUSE Linux Enterprise 9
DSpace	CONTENTdm and DigitalCommons (vender-hosted)	Dell PowerEdge 2850	4 GB	3 TB	Red Hat Linux 3
	CONTENTdm	Dell PowerEdge 2600	2 GB	360 GB	Windows 2000 Server
	CONTENTdm	Dell PowerEdge 2650	2 GB	2 TB	Windows 2000 Server
		Intel, Sun			Windows, Sun
		SAN	4 GB	600 GB	Red Hat Linux

Planning an IR N=20

Freeware	Commercial Software	Computer Model	Memory	Storage	Operating System
DSpace (4)					
DSpace					Linux
DSpace					Solaris
DSpace		Dell PowerEdge 2650	512 MB	108 GB	SUSE Linux
DSpace		Dell PowerEdge 6650	8 GB	932 GB	Linux

DSpace		HP Proliant DL385	12 GB	3 TB	Red Hat Linux Enterprise AS3
DSpace		IBM RS/6000	4 GB	275 GB	IBM AIX 5L
DSpace		Sun (three machines)			
DSpace	Digitool				
DSpace, Fedora					Linux
DSpace, Fedora	Aware image processing software	Dell 2850's	4 GB	3 TB	Windows 2003 server/ clustering server
Fedora					
Fedora					Red Hat Linux
Greenstone		Dell	4 GB	500 GB	Linux
	CONTENTdm	Dell PowerEdge 350, Dell AX-100, Gateway 9415		2.8 TB	Linux, Windows
	Digitool	IBM/Linux/Intel		TB	Linux
	Documentum				

Vendor-hosted IR N=7

ProQuest Digital Commons	6
Bepress tools	1

11. Has the IR software been modified to enhance its functionality (not just to "brand" the interface)? N=52

	Total N=52		Have N=36		Planning N=16	
A few minor changes have been made	17	33%	12	33%	5	31%
No changes have been made	16	31%	10	28%	6	38%
We have made major modifications to the software	10	19%	6	17%	4	25%
We have made frequent changes on an ongoing basis	9	17%	8	22%	1	6%

Selected Comments from Respondents

Minor Changes

"We created a quick submit button to allow submitters to enter a limited amount of metadata."

"Modifications have been made to allow for LDAP authentication of users."

"Metadata changes and search capabilities."

"Works with campus authentication system"

"Bitstream modifications. Local authentication and authorization."

"RSS feeds"

"A few minor changes have been made (the University of Alberta has adapted the University of Manitoba code to generate Library and Archive Canada's ETD metadata)."

"'Researcher Pages' added to current release. Work on authentication to local system."

"Plan to perform greater customization using the Manakin XML UI for DSpace."

Major Modifications

"CDL's version of the bepress software is highly modified from the original EdiKit, though it is not very different from that available through Digital Commons, i.e., most of the co-developed changes have been carried into DC."

"Our IR consists of multiple systems running a variety of IR software. These expose their metadata to a metadata repository that provides a harvester and user interface for searching and linking to them, and we built these pieces ourselves. To access grid resources (e.g., research datasets) that we represent in our IR, we developed an OAI-PMH-to-SRB translator called OAISRB. User interfaces have been customized to represent a consistent look-and-feel with other library systems."

"Changed the complete interface system."

"Out of the box, Greenstone 'holds' the metadata and the digital resources/objects, allows them to be searched, and creates results lists. The default Greenstone presentation layer (display on the Web) is not very customizable so MUCH work has been done in house to create a new/different/custom presentation layer."

"We have been heavily customizing the workflow of DSpace and have made other customizations. We are also actively contributing to the DSpace development community."

Frequent Changes

"Metadata—ETD-ms, Scorm Crosswalk to MARC. Multiple submission types created. Help Screens modified and created. Interface modified."

"Minor layout adjustments, some operational modifications."

POLICIES AND PROCEDURES

12. Does the IR have written policies and procedures? N=57

	Total N=57		Have N=36		Planning N=21	
Yes	42	74%	27	75%	15	71%
No	15	26%	9	25%	6	29%

If yes, have IR policies and procedures been reviewed by institutional officials, such as university counsel? N=41

	Total N=41		Have N=28		Planning N=13	
Yes	22	54%	15	54%	7	54%
No	19	46%	13	46%	6	46%

If yes, please identify the title of the reviewer(s). N=14

University Counsel (10)

Vice-Provost (2)

Dean of Libraries (2)

Attorney General

Faculty Senate

Computer Science faculty

University Provost

University Task Force on Digitization Initiatives (Chaired by Director of Libraries)

Graduate Dean

Graduate and Research Council

Advisory Committee

13. What types of materials are included in your IR? Check all that apply. N=56

	Total N=59		Have N=36		Planning N=19	
Electronic theses and dissertations	39	70%	24	67%	15	79%
Articles, preprints	35	63%	22	61%	13	68%
Articles, postprints (author modifies preprint to match published work	34	61%	22	61%	12	63%
Conference presentations	32	57%	18	50%	14	74%
Technical reports	31	55%	18	50%	13	68%
Working papers	31	55%	19	53%	12	63%
Articles, official published version	29	52%	16	44%	13	68%
Conference proceedings	26	46%	15	42%	11	58%
Multimedia materials (locally created for scholarly purposes)	25	45%	15	42%	10	53%
Electronic journals (locally published)	24	43%	15	42%	9	47%
Student-produced documents	23	41%	14	39%	9	47%
Datasets	20	36%	10	28%	10	53%
Learning objects	19	34%	10	28%	9	47%
Books, official published version	15	27%	10	28%	5	26%
Books, author draft or author modified draft to match published work	12	21%	6	17%	6	32%
Institutional or unit annual reports	12	21%	4	11%	8	42%
Institutional magazines or newspapers	11	20%	5	14%	6	32%
Working documents of administrative/support units	9	16%	4	11%	5	26%
University catalogs	5	9%	1	3%	4	21%
Overlay journals	4	7%	—	—	4	21%
Alumni association publications	4	7%	—	—	4	21%
Yearbooks	3	5%	—	—	3	16%
Other type of material	25	45%	14	39%	11	58%

Please specify other type of material.

Have an IR

Newsletters (2)

Edited volumes; monographic series

Videotaped conferences & presentations

Color digital images of human tissue arrays, histology images, historic maps

Practica

Book chapters

Electronic monographs and reference works (no print editions)

Texts of lecture series

Data sets

Musical scores; historical photographs

Digital images; audio/video files (e.g. oral history); archival manuscripts

Digital photographs

Research project(s) from academic units

Planning an IR

Agricultural training document (EDIS)

Landmark documents in university history

Locally published lecture and monograph series

Digitized special collections; image database

Archival research resources

"The pilot project will collect electronic theses; other materials will be considered for addition at a later date."

"All formats are under consideration. Special collections in digital format will be emphasized."

14. What types of digital files (e.g., Acrobat, ASCII, HTML, PostScript, PowerPoint, TIFF, Word, and XML) are acceptable for the IR? (NB: Preservation implies that data will not only continue to be maintained, but also continue to be made accessible, using data migration to new digital formats or other techniques to ensure access in a changing technological environment.) N=52

	Total N=52		Have N=34		Planning N=18	
Any digital file type is accepted, but only some types are preserved	24	46%	16	47%	8	44%
Any digital file type is accepted and preserved	16	31%	9	26%	7	39%
Only specified digital file types are accepted and preserved	9	17%	6	18%	3	17%
Only specified digital file types are accepted (there is no preservation)	3	6%	3	9%	—	—

Please specify file type.

Any type accepted, some preserved

XML, ASCII, PDF (migration to PDF/A when available), TIFF, JPEG2000

PDF, AIFF, AIF, AIFC, GIF, HTML, HTM, JPEG, JPG, MARC, PNG, PS, ES, AI, RTF, TXT, TIFF, TIF, XML

PDF, JPEG, PNG, GIF, TIFF, AIFF, MPEG, XML

ASCII, XML, Acrobat

XML, TIFF, TXT, RFT, PS, EPS, AI, PNG, MARC, JPG, HTM, GIF, AIFF, PDF

PDF, Word, many others

ASCII

PDF, HTML, TIFF, XML, JPEG

AIFF, AVE, JFIF, JP2, JPX, PDF, TIFF, WAVE, XML

"We follow the DSpace preservation categories."

"Same as DSpace."

"Standards-based formats such as TIFF, PDF/A."

"Those not proprietary."

Specified types accepted and preserved

Word processing files, RTF, PDF (all files converted to PDF)

PDF, TIFF

PDF

MARC, Adobe PDF and Postscript, TIPP, GIF, PNG, HTML, TXT, DAT, ASCII, RTF XML, AIFF, AIF, AIFC

PDF, HTML, ASCII, XML, SGML, JPG, PNG, TIFF, GIF, CGM, PPD, EPS, MPEG-2, MP-3, XCL, DXF, GIS

TIFF, PDF, HTML, WAV/AIF

Specified types accepted, none preserved

PDF

PDF (micro-filmed for preservation), TIFF, GIF, JPEG, AMOV, WMV, AVI, RM, MPEG, AIF, MIDI, SND, WAV, SWF

"PDF for main work—any digital file type for associated work."

15. What is the document deposit procedure for the IR? Check all that apply. N=55

	Total N=55		Have N=36		Planning N=19	
Authorized depositors deposit documents directly	47	94%	32	89%	15	79%
IR staff deposit documents for authorized users	40	73%	28	78%	12	63%
Other procedure	11	20%	5	14%	6	32%

Please describe other procedure.

"Academic units assign administrative staff to upload objects."

"Administrative staff submit documents for authorized users."

"For now, IR staff only; in future, IR staff AND authorized depositors can both submit."

"In pilot, IR Staff will deposit docs."

"Input workflows are under construction; manual deposit will be secondary."

"Libraries staff deposit theses and scanned back-files."

"Library staff (not connected to the IR) deposit documents."

"Procedure to be finalized."

"Proxy depositing by the administrative staff of departments."

"Unique, non-standard documents require special consultations with consortium."

"We have a few who deposit directly, but it's not the norm."

16. Whose materials may be deposited in the IR? Check all that apply. N=55

	Total N=55		Have N=36		Planning N=19	
Faculty	52	95%	34	94%	18	95%
Non-faculty professional staff	40	73%	26	72%	14	74%
Students	40	73%	25	69%	15	79%
Support staff	23	42%	13	36%	10	53%
Other category	19	35%	12	33%	7	37%

Please specify other depositor category.

"Affiliated hospital consortium"

"Affiliated staff (State Surveys located on campus)"

"All contributors must be approved by IR team"

"Any faculty-sponsored materials"

"Any materials must have approval of the department chair or research center director."

"Any UC author of published material can deposit a postprint."

"Anything that is faculty sponsored, so it could include student papers."

"Document produced by organizations, such as Chemistry Department newsletter and financial reports of university administration"

"Faculty or departments may sponsor student work for deposit."

"Librarians"

"Library archival collections"

"Masters/doctoral theses will be accepted during the pilot; range of materials will increase over time."

"University press books"

"Outside faculty and contributors"

"Product from digitization initiatives"

"Research partners external to the university"

"Special collections in digital format"

"Too early to tell whose materials will be deposited."

17. **Are documents that are submitted by authorized users reviewed and approved for deposit to ensure copyright compliance or for other reasons? N=49**

	Total N=49		Have N=34		Planning N=15	
Documents are not reviewed (authors are solely responsible for their documents)	18	37%	14	41%	4	27%
Designated departmental or unit officials review and approve documents	16	33%	12	35%	4	27%
IR staff review and approve documents	9	18%	7	21%	2	13%
Other reviewer	6	12%	1	3%	5	33%

Please describe other reviewer.

Selected Comments from Respondents

"Some items are reviewed by departmental or unit officials, others are not reviewed at all."

"For materials deposited by library staff, library staff review for copyright compliance."

"Each collection has a different policy for whether submissions are reviewed, by whom, and for what purpose. It's delegated at the point when a new community is established."

"Policy is under revision. There is a review process but currently not well followed."

18. Do authorized depositors sign a deposit agreement (including click-though Web form agreements)? N=50

	Total N=50		Have N=35		Planning N=15	
Yes	44	88%	31	89%	13	87%
No	6	12%	4	11%	2	13%

19. Does your library negotiate with publishers to allow for the permanent deposit of e-prints from their serials in the IR? N=51

	Total N=51		Have N=35		Planning N=16	
No, but such negotiations are being considered	26	51%	16	46%	10	63%
No, and there are no plans to do so	19	37%	14	40%	5	31%
Yes	6	12%	5	14%	1	6%

20. Are multiple versions of the same document permitted? N=48

	Total N=48		Have N=35		Planning N=13	
Yes	32	67%	24	69%	8	62%
No	16	33%	11	31%	5	38%

If yes, please describe techniques to control versioning.

Selected Comments from Respondents

"Superceding works are deposited and linked to original record; a link to the previous version is presented on the metadata page."

"Multiple versions are held within the system, but only one version is viewable at a time."

"DSpace permits versioning."

"Author's control."

"We add the revised document to the original digital object, with a date and additional metadata describing the revision."

"It is allowed; we have accepted multiple versions and have noted this in the subsequent records. We are looking at modifying the software code to do this better."

"Submitter can: note version in document title; include information in relation element of metadata record; use date of issue to indicate version."

"Communities within the IR have control over what they allow for versioning. They are encouraged, though, to submit last version only."

"In some series, older versions are archived."

"No specific versioning provisions beyond submit date."

"We have this capability but haven't used it as yet."

"Software and metadata support version control."

"At present, there is no special control."

"Multiple versions are one item in DSpace with multiple files."

"Included within the same deposit, but the file names indicate the different versions. One version does NOT replace another."

"Pre-prints may be withdrawn and replaced with post prints. For other materials, multiple versions, showing progression are preferred."

"Versioning is done merely by adding dates of creation to the metadata."

"CVS style versioning native to the repository software."

"Author's discretion."

"Qualified DC allows for version info."

"Potentially a pre- and a post-print might be included in the repository. We will use date published and inclusion of citation information to indicate versions."

"Currently in pilot planning phase; techniques for versioning control unknown at this time."

"Will have to be encoded in work practices (given the current functionality of the DSpace software)."

21. Can documents be withdrawn from the IR? N=45

	Total N=45		Have N=34		Planning N=11	
Documents can be withdrawn	37	82%	28	82%	9	82%
Documents cannot be withdrawn	8	18%	6	18%	2	18%

If documents can be withdrawn, who has authority to do so? Check all that apply. N=39

	Total N=39		Have N=29		Planning N=10	
IR staff can withdraw documents only under specific circumstances	26	72%	20	69%	6	60%
Authors can withdraw documents only under specific circumstances	9	25%	7	24%	2	20%
IR staff can withdraw documents at will	8	22%	5	17%	3	30%
Authors can withdraw documents at will	6	15%	4	14%	2	20%
Other authority	4	11%	1	3%	3	30%

Selected Comments from Respondents

IR staff, specific circumstances

"Only if there are copyright problems."

"A 'tombstone' will remain in system with a message about withdrawal."

"Copyright infringement, fraud etc."

"Copyright violation, content not within policy guidelines."

"Only non-peer reviewed works may be withdrawn (they are actually hidden; the metadata is not, so no dead links). Peer-reviewed works, like books and articles may not be withdrawn."

"At the author's request."

"Issues of copyright, plagiarism, falsified data."

"In the case of fraud, serious factual errors, or as a legal requirement."

"At request of author; at the direction of the Provost; by legal order."

"If requested by authorized departmental depositor."

"Infringement of copyright (this hasn't happened yet)."

"Author request."

"Unintended duplication, copyright violation, remain flexible for ad hoc circumstances."

"In consultation with appropriate parties."

"Licensing, rights issues."

"If document was mistakenly entered or there is an issue with copyright or appropriateness."

"Legal issue or to protect intellectual property."

"In cases of legal necessity; metadata for withdrawn items are retained."

"As copyright becomes more of an issue, we require withdrawal flexibility at local university level."

"Upon request of author."

"By author request or for legal reasons."

"Needs to be encoded in a policy; one example would be research misconduct."

Author, specific circumstances

"Only if there are copyright problems."

"Their rational may not be based on a change of heart about the content; only under legal threat from outside or plagiarism, will we remove items."

"A 'tombstone' will remain in system with a message about withdrawal."

"Preprints can be replaced with postprints."

"Tombstone or reference of removed documents remains."

"Authors can request withdrawal, and reasons/terms for this are negotiated on a case-by-case basis."

"Guidelines are determined by each community."

Other authority

"Advisory committee retains the right to withdraw materials in numerous situations."

"All withdrawn materials will be traced in the form of a note field in the Dublin Core record. We will supply a tombstone when the withdrawn item is requested. Tombstone will include original metadata and withdrawal note."

"Again, as with multiple versioning, this option is still in discussion."

22. Does the IR import metadata from external systems? N=51

	Total N=51		Have N=36		Planning N=15	
No	26	51%	19	53%	7	46%
Yes, external metadata are mapped to IR metadata format(s) by system programs	19	37%	13	36%	6	40%
Yes, external metadata are mapped to IR metadata format(s) by hand	5	10%	4	11%	1	7%
Yes, the external metadata are in the same format(s) as internal IR metadata	1	2%	—	—	1	7%

List format(s) of external/internal metadata.

EAD/EAD2002; OAI/DC; XML-MARC; FGDC; RSS

Briefly describe the types of metadata and mapping by hand.

"Dublin Core"

"For e-theses cataloging staff add LCSH headings."

"Locally created metadata schemes are mapped to DSpace qualified Dublin Core."

Briefly describe the types of metadata and mapping by system programs.

"Web of Science citations -> eScholarship Repository postprint MD schema"

"XML metadata maps to DC and MARC"

"Initially items are mapped by hand, then process is automated and mapped by programs."

"A crosswalk from MARC to Modified Dublin Core"

"Excel tables (CVS mapping)"

"ETD-db tables for electronic theses and dissertations; custom Oracle database for faculty publications and bibliographic data"

"DC, MARC, FGDC/DIF, custom"

"Batch import scripts from Excel spreadsheets into DSpace"

"MARC metadata is transformed into Dublin Core"

"Metadata scraped from Web sites listing publications; bibliographic data is also imported into and from bib

databases."

"PubMed"

"MARC to DC"

"For titles cataloged in our OPAC (the ExLibris product, Aleph), we have a map from MARC 21 to our internal METS schema, which we call UFDC Mets. For author self submittals, we currently plan to use a Web form, which will use basic pieces from the UFDC mets, with a touch of human intervention."

"Dublin Core"

"Local formats in department or library databases; initial mapping is done by hand but will be automated once the initial mapping is done."

"Under development along with anticipated implementation of CONTENTdm."

"MARC to Dublin Core"

23. Who enters metadata for deposited documents? Check all that apply. N=53

	Total N=53		Have N=35		Planning N=18	
Authorized depositors enter simple metadata (e.g., Dublin core)	47	89%	33	94%	14	78%
IR staff enhance depositor-supplied metadata	33	62%	20	57%	13	72%
IR staff enter simple metadata for authorized users	31	59%	21	60%	10	56%
IR staff catalog material completely based on local standards	14	26%	11	31%	3	17%
Other procedure	11	21%	5	14%	6	33%

Please describe other procedure.

Selected Comments from Respondents

"A program in the system creates and enters the metadata automatically."

"Cataloging staff provide metadata for local material; also some is handled in batches from other databases and library catalog where applicable."

"Currently in pilot planning phase; to be decided."

"If material is in local catalog (ExLibris OPAC), metadata can be imported in via Excel batch loader."

"IR staff will apply national standards as available."

"Metadata entered through independent interface by depositors."

"Mixed—IR staff & depositors."

"Postprints MD ingested from harvested citations (after being mailed to faculty)."

"We are investigating including user (i.e., searchers/browsers of the IR) supplied metadata but are not doing that yet."

24. **What efforts have been made to ensure or enhance IR interoperability with external systems? Check all that apply. N=52**

	Total N=52		Have N=35		Planning N=17	
The IR supports OAI-PMH	48	92%	33	94%	15	88%
Persistent identifiers are used for IR materials	45	87%	31	89%	14	82%
The IR is OpenURL compliant	26	50%	19	54%	7	41%
IR content is included in federated searching systems	23	44%	22	63%	1	6%
Other effort	7	14%	3	9%	4	24%

Please describe other effort.

Selected Comments from Respondents

"Crawled by Google and Google Scholar."

"Currently in pilot planning phase; other efforts to be decided."

"Interoperability evaluation & analysis through grant funding."

"The eScholarship Repository is an XML Gateway."

"Theses Canada Portal and NDLTD union catalogue harvest metadata."

"University System developed shared metadata guidelines."

CONTENT RECRUITMENT

25. **What strategies have been used to recruit content? Check all that apply. N=57**

	Total N=57		Have N=36		Planning N=21	
IR staff have actively identified likely depositors (e.g., users who have already done self-archiving, authors who publish in open access journals, or authors who publish in journals with policies that support self-archiving) and encouraged them to submit materials	43	75%	28	78%	15	71%
IR staff have made presentations to faculty and others	42	74%	30	83%	12	57%

Subject specialists in the library have acted as advocates with their faculty	40	70%	28	78%	12	57%
IR staff have offered to deposit electronic materials for authors	34	60%	23	64%	11	52%
IR staff have offered to digitize printed materials and deposit them for authors	27	47%	18	50%	9	43%
Faculty and/or administrators have encouraged deposit by others	25	44%	21	58%	4	19%
Electronic announcements via e-mail, the Web, or other means	25	44%	19	53%	6	29%
Symposiums and conferences have been held to raise awareness of related scholarly communication issues	20	35%	14	39%	6	29%
Other strategy	19	33%	13	36%	6	29%

Please describe other strategy.

Selected Comments from Respondents

"Brochure Web sites; flyers; PPT available on Web site."

"No promoting."

"We've contracted with the campus student newspaper to get their stories."

"Upcoming visit to campus by Cliff Lynch."

"Getting contributors to advocate with spouses and friends."

"Setting up publication alerts on relevant databases, then asking faculty for that content in a timely manner."

"We have begun with electronic theses and dissertations. We are currently developing an approach to content recruitment."

"Brochures, article in newsletter."

"IR staff have made presentations to graduate students to publicize and explain the electronic thesis submission option."

"Library staff have begun harvesting materials from Web sites in particular colleges."

"Collaboration with active sponsored research projects."

"IR staff examine Web sites on campus and gather materials for inclusion."

"Incorporated existing digital collections, such as ETDs and accreditation documents database."

"Contacts with Library Digital Program and Center for Educational Resources staff."

26. Is there any pressure on authors to submit content to the IR? N=52

	Total N=52		Have N=36		Planning N=16	
No pressure on authors to submit content	28	54%	18	50%	10	63%
They are encouraged to do so	21	40%	16	44%	5	31%
They are not required to do so, but this is being actively considered	2	4%	1	3%	1	6%
They are required to do so	1	2%	1	3%	—	—

If required, please explain.

"Students must submit theses or dissertations."

27. On a scale of 1 to 5—where 1 is very easy and 5 is very difficult—please indicate how much effort has been required to obtain materials to deposit in the IR. N=48

	N	Very easy	Somewhat easy	Neutral	Somewhat difficult	Very difficult
Total	48	4	7	12	15	10
Have	35	1	7	5	14	8
Planning	13	3	—	7	1	2

Please explain.

Selected Comments from Respondents

Very easy

"They have come to us."

"As word has spread about our collaborations with research investigators, others come forward seeking similar solutions to data management/cyberinfrastructure issues."

"At this time, there are many, many digital publications available on the Web site that are possible IR content. In an early phase, these materials will be harvested by IR staff. Over time, wider sets of intellectual content will be recruited from faculty, graduate students, etc."

"It has been very easy to identify content for the pilot. We anticipate high interest in specific arenas when we move to production mode."

Somewhat easy

"It has been easier than we expected to get the first 10,000 papers deposited by faculty departments into the repository, but what it will take to attract a significant percentage of the 28,000 UC authored articles

is unknown, but likely to be nothing short of a mandate. At this time, the system-wide academic council is considering a proposal that UC faculty routinely grant to UC a non-exclusive license to manage their content in the institutional repository."

"We work with the Graduate College to get a copy of the electronic theses or dissertation and the distribution rights from students. Students are required to submit the ETD to the Graduate College but could deny distribution rights."

"Faculty have been very receptive to offering materials for inclusion."

"We have identified candidate collections from on-campus sources (library and academic) that would most likely collaborate in the IR development and whose content would be of greatest benefit to our community. We work continuously to build collaborations among these units and to resource collection development toward shared goals."

"The percentage of theses submitted electronically has been increasing every year since it became an option in 1999."

Neutral

"Still in pilot state of the IR. Need easier to use interfaces before an aggressive campaign is launched to recruit contributions."

"We have not put any effort into this."

"We're really just starting to figure this out. Some have been really easy, others more difficult."

"Currently, it's again a question of policy—the academic affairs side of the university has a very diverse body of departments and schools and every single one is ripe for inclusion in the IR. But the framework for effective IR policy at the local university level is not yet created. Spring 2006 much work is planned to bolster the IR operations, both via the consortium model as well as from central university information technology services."

"We have not yet expended much effort to recruit content as we are still in the implementation phase. We have identified a set of early adopters but have not moved beyond this group."

"University administrators and faculty, as a whole, are interested and enthusiastic. Individual units sometimes express reluctance due to perceived loss of control."

Somewhat difficult

"Faculty operate very independently and think they can do this on their own."

"We have one collection where there seems to be a high motivation to do submissions, but it is too early to say if that will pan out. Apart from that, our repository has very few items in it to date. Having said that, it is largely due to the fact that we have not systematically promoted it."

"Working with the Faculty of Graduate Studies and Learning Technology Centre has been easy because they could see we were solving a problem for them. Approaching other faculties has been more difficult because the reason for the IR is not perceived as clearly."

"With more than 50 communities in Dspace there is no single answer to this. Some departments have been

supporters from the beginning; others have been very reluctant to participate. Within departments some faculty are very active depositors, others deposit nothing. Recruitment of some 'types' of content is simple, other types more difficult (especially published articles)."

"Our IR is still young. We haven't reached a momentum yet that would allow us to recruit content more easily."

"Not a priority for most faculty."

"There is a reluctance on the part of faculty and research community to participate."

"Initially, it was very difficult. Now that we have acquired a critical mass, more departments and individual authors are approaching us. It is getting easier to acquire new content every day."

"Recruiting and managing copyrighted material—a core collecting goal—has been difficult and time-consuming."

"A marketing team is working on motivation and incentive programming; currently there is a focus on working with faculty data sets."

"Faculty need regular encouragement and reminders."

"Difficult concept to explain. Faculty often prefer their own sites or want to retain the ability to remove materials at will. Not as compelling a task as research and teaching. Copyright difficult. Overall too hard to give up control and add to workload."

Very difficult
"The problem is more related to a lack of marketing of the IR."

"Faculty are reluctant to deposit for a variety of reasons—too hard, forgetfulness, don't see the need, etc."

"Authors don't see any advantage yet to do so. We are planning to input, soon, a critical mass of content with the ETD and to implement a version 2 of Archimede in order to augment the visibility of our IR."

"Faculty are very support and even enthusiastic about the IR when it is presented to them but few actually deposit or if they do, few persist in depositing."

"Faculty who are approached are supportive of the idea, but usually too busy to submit content. Copyright is often a concern as well."

"IF we do the work, most will give us documents, but authors are generally unwilling to go through the multi-step process themselves."

"We are only at beginning stages and have not yet embarked on campus-wide content recruitment phase."

"From reading anecdotal experiences of others and investigating existing IRs, it is clear that populating IRs is a significant challenge."

ASSESSMENT

28. Has your institution done research on why users contribute or do not contribute documents to the IR? N=54

	Total N=54		Have N=36		Planning N=18	
Yes	13	24%	8	22%	5	28%
No	41	76%	28	78%	13	72%

If yes, please describe the research.

Selected Comments from Respondents

"We are launching a survey next week on faculty attitudes toward their IP/rights management and alternative publishing mechanisms like the eScholarship Repository. Stay tuned."

"Conducted faculty focus groups."

"There was a nine-month project to investigate recruitment strategies using three sample departments and looking at factors such as copyright clearance for older material, support for scanning and metadata entry, and various marketing strategies."

"Literature search; networking with similar institutions; discussions with others at conferences, etc."

"No research on [local] contributions, but have carefully followed other published research accounts of IR adoption."

"Series of recurrent interviews with early adopters. Interviews with key respondents from within the library. Survey of users and reference librarians."

"We conducted an internal survey to find out why librarians won't submit to the IR. Reasons were: no time and not sure their materials are worth preserving."

"Readings in professional literature."

"Read articles. Attended many conference sessions."

"Research to date has consisted of reviewing literature. A formalized investigation will be conducted at a later stage in our pilot phase."

"Interviews of potential users. (IR is not implemented)."

29. What techniques have been used to evaluate the success of the IR? Check all that apply. N=35

	Total N=35		Have N=29		Planning N=6	
Tracking hits on IR content	24	69%	23	79%	1	17%
Interview external users (e.g., researchers, faculty, students) of the IR	11	31%	9	31%	2	33%
Interview internal users (e.g., library staff) of the IR	8	23%	7	24%	1	17%
Citation analysis on IR content	5	14%	4	14%	1	17%
Conduct user surveys	4	11%	3	10%	1	17%
Conduct focus groups	3	9%	1	3%	2	33%
Other technique	10	29%	7	24%	3	50%

Please specify other technique.

Selected Comments from Respondents

"Currently designing a usability study; informal collaborator/user surveys."

"Currently in pilot planning phase; evaluation methods to be decided at later date."

"Level of participation/contribution."

"Not evaluated yet. Waiting for version 2 + critical mass of content."

"Track collection growth to identify key adopters, and monitor conditions of the vendor contract."

"We are in the process of developing evaluation techniques."

"We count actual full-text downloads rather than hits, and the statistics are distributed monthly to authors. They in turn provide a lot of feedback about the increased visibility and use that their work receives in the eScholarship Repository."

CURRENT STATUS OF IR

30. How many digital objects (e.g., digital audio files, images, videos, technical reports, e-prints, theses, dissertations, etc.) are currently deposited in the IR? N=39

	N	Minimum	Maximum	Mean	Median	Std Dev
Total	39	4	19,170	3,479.5	1,996.0	4,390.4
Have	34	20	19,170	3,844.5	2,211.5	4,545.1
Planning	5	4	4,500	997.0	106.0	1,963.1

	Total	Have	Planning
<100	7	5	2
100–999	10	8	2
1,000–9,999	17	16	1
>10,000	5	5	—

31. How much disk storage is required to house existing materials? N=28

	Total	Have	Planning
<1 GB	5	5	—
1–9 GB	9	7	2
10–99 GB	7	6	1
100–999 GB	4	3	1
≥1 TB	3	3	—

32. What is the typical number of end-users at peak times? N=38

Most respondents were unable to answer this question. Of those who track usage, responses for number of end-users ranged from 1 to 1100 at peak times to 500 per day to 55,000 per week. Those who track the number of downloads reported from 60 to 300 per day.

33. Are parts of your IR restricted to specific user groups (e.g., documents produced by a department are only deposited and accessed by department members)? N=50

	Total N=50		Have N=36		Planning N=14	
Yes	21	42%	16	44%	5	36%
No	29	58%	20	56%	9	64%

If yes, please describe the restrictions.

Selected Comments from Respondents

"Dissertations are currently available only through dissertation distributor's subscription."

"We have a closed collection for theses and dissertations to protect students with pending patents or copyright issues."

"Retrospective dissertations are for campus use only. Also a few miscellaneous items have been handled this way due to potential legal constraints."

"Not currently, but we will eventually have collections that will be restricted to an individual college on campus."

"For ETDs—Restricted collection available to students requesting this option."

"Collections can be limited to the MIT community at the discretion of the owner."

"Some department materials are restricted to faculty in that department. E-dissertations are available but not free to non-university users."

"Access to dissertation restricted to university affiliates."

"IP, user account"

"Certain faculty collections have requested restricted access due to copyright issues with images; cultural concerns with primary source materials from tribal communities."

"Some have campus only restrictions."

"Images derived from copyrighted materials (i.e., copystand photography) are restricted to IP_authenticated university users and affiliates."

"Some material is limited to campus users only. Some is limited to the members of the depositing department or to a list of authorized users. Authentication is currently being done with default DSpace user management, but will be set to use campus authentication system (EIDs)."

"Authors are free to restrict access to institutional IPs (all, not specific departments.) We instituted this policy, reluctantly, after user feedback. To date, no one has used it."

"One community has images which are currently available only to those who are members of an international research consortium."

"The research project material from one academic unit is restricted to project members."

"Subject to copyright and individual author/departmental policy."

"We have not yet implemented restrictions but will allow access restrictions on a limited basis to either a group of identified users or to only the campus."

"Pilot will help us determine levels of security required."

"Currently in pilot planning phase; plan to allow for some access controls, as needed, in pilot project."

34. Are any IR documents supplied to external users for commercial purposes with associated use fees going to the institution? N=49

	Total N=49		Have N=35		Planning N=14	
Yes	4	8%	3	9%	1	7%
No	45	92%	32	91%	13	93%

If yes, please describe the commercial use.

Selected Comments from Respondents

"ETDs— ProQuest"

"Image re-use is granted per request; fees are negotiated by the collection owner."

"Institutional images."

"When the photographic collections of historic images become accessible via Content DM, commercial users such as publishers and broadcast media will continue to pay fees for high resolution or darkroom images plus fees for publication and other use."

BENEFITS AND CHALLENGES

35. What are the top three benefits you feel your IR provides? N=50

Enhance visibility and increase dissemination of institution's scholarship	34	68%
Free, open, timely access to scholarship	23	46%
Preservation of and long-term access to institution's scholarship	18	36%
Preservation and stewardship of digital content	18	36%
Collecting, organizing assets in a central location	12	24%
Educate faculty about copyright, open access, scholarly communication	4	8%

36. What are the top three challenges that your institution has faced in implementing, promoting, and running the IR? N=50

Content recruitment, building a critical mass of content	16	32%
Staffing issues	15	30%
Faculty awareness/buy-in/interest/engagement	14	28%
Copyright issues	11	22%
Communicating with faculty, articulating the benefits of the IR	10	20%
Adequate funding and other resources	7	14%
Integrating a new unit/workflow into existing structures	6	12%

ADDITIONAL COMMENTS

37. Please submit any additional information regarding the institutional repository at your institution that may assist us in accurately analyzing the results of this survey. N=26

Selected Comments from Respondents

"Additional challenges include getting other Internet sites to link, search, and recognize content; getting scholars and administrators to accept e-publication as a viable credential."

"Although we have used DSpace for about two years, I would not say that we have an active IR. In fact we have little more than a pilot project at this point. Our locally published scholarly journals are found in CONTENTdm, as well as our ETD collection."

"Another challenge: providing ongoing staff support, ensuring sustainability."

"Dspace@MIT is somewhat unusual among production IRs because of its age (>3 years in production) and having been developed in-house with HP Labs."

"[Our] use of the consortium IR is a small movement in the establishment of an overall university-wide approach to digital asset management. Due to the culture of distributed computing, it has been extremely difficult to both advocate for a good strategy for an IR as well as articulate to the campus stakeholders why an IR is of use. Thanks to the advances of the consortium implementation, the university has at least an opportunity to understand how IR's can be used on campus.

In Dec 2005, I made a detailed presentation on digital libraries and their future to a campus wide Research and Information Technology Committee. The consortium IR was featured at the end as a significant opportunity for the university to embrace new ways to deal with campus-generated content. In Summer 2005, I was honored to be an Educause Frye Institute participant and part of my acceptance into the institute was based on my proposed project towards creating a true institutional repository. Situations on campus have begun to

shift and the opportunities to bring disparate groups together to benefit from *enterprise* wide planning, implementation, and long-term maintenance finally seem to be appearing. I can honestly say there is a wealth of opportunity at the campus level—but it is only going to be effective if the library can 'sell' the idea at the top levels of university administration."

"I have described the content of one 'flavor' of a digital repository at our institution. Other content, for example, digital dissertations, are contained within a separate infrastructure. The amalgam of several structures under one IR umbrella is currently being discussed and will more fully comprise a holistic IR for our campus."

"We currently have 8200+ ETDs in our IR. We will be attempting to capture all Honors College Senior theses for those graduating in Spring 2006."

"Our next effort at populating the IR will be to work our 71 Regents Professors to showcase their scholarly work in conjunction with a photo gallery display of Regents Profs in the Library."

"IR development is in a preliminary stage so many questions have been left blank. A pilot project to seek submissions of electronic theses and dissertations is currently underway."

"The IR is not yet in production. A task force of librarians and central IT staff has been formed. To date, the task force has done some preliminary investigation and hosted a visit by Cliff Lynch with attendees from each of the colleges. The task force is currently examining and documenting the University's entire information landscape in order to determine next steps."

"IR is centrally run for the benefit of 10 campuses."

"Helping to lead planning process for a University System-wide repository including metasearching of contents and IR hosting services."

"The committee charged with developing a pilot IR formed in September 2005. It is anticipated that the pilot project will go online in early to mid-2007; a full IR roll-out is anticipated for 2008/09."

"The resource is in development with proposed rollout Jan 2007. We do not use the expression 'institutional repository' as this expression does not convey significant meaning; indeed, it does not convey a message of success."

"The system we are building is distributed institutional repository, which focuses currently on investigating and resolving issues related to data sets (collecting, describing, curating, archiving, etc.). Documentation is being drafted, is under review, and not available at the time of this survey."

"There are no immediate plans to develop an IR, but this will be re-assessed in six months."

"This is still a work in progress. We are planning the launch of version 2 in March 2006. It will be a major upgrade, able to accommodate a large number of DTD and schemas and we hope that this version will stimulate more interest."

"We are really struggling to get the IR off the ground and the lack of human resources dedicated to the initiative has been the major barrier."

"We found this survey to be very helpful in terms of giving us ideas on how to move forward (for example, what specific kinds of information could be available on a public Web site; types of content to suggest to potential submitters for the IR system.)"

RESPONDING INSTITUTIONS

University of Alabama
University at Albany, SUNY
University of Alberta
University of Arizona
Arizona State University
Auburn University
Boston College
Brigham Young University
University of British Columbia
University at Buffalo, SUNY
University of California, California Digital Library
University of California, Davis
University of California, Irvine
University of California, Los Angeles
University of California, San Diego
University of California, Santa Barbara
Canada Institute for Scientific and Technical Information
Case Western Reserve University
University of Chicago
University of Colorado at Boulder
University of Connecticut
Cornell University
University of Delaware
University of Florida
George Washington University
University of Georgia
Georgia Institute of Technology
University of Guelph
University of Hawaii at Manoa
University of Houston
University of Illinois at Chicago
University of Illinois at Urbana-Champaign
Indiana University Bloomington
University of Iowa
Iowa State University
Johns Hopkins University
University of Kansas
Kent State University
University of Kentucky
Université Laval
Library of Congress
Louisiana State University
University of Louisville
McGill University

McMaster University
University of Manitoba
University of Massachusetts Amherst
Massachusetts Institute of Technology
University of Michigan
University of Minnesota
Université de Montréal
University of Nebraska–Lincoln
New York University
University of North Carolina at Chapel Hill
North Carolina State University
Northwestern University
University of Notre Dame
Ohio University
Ohio State University
University of Oklahoma
Oklahoma State University
University of Oregon
University of Pennsylvania
Pennsylvania State University
University of Pittsburgh
Purdue University
Queen's University
Rice University
University of Rochester
University of Southern California
Southern Illinois University Carbondale
University of Tennessee
University of Texas at Austin
Texas A&M University
University of Toronto
University of Utah
Vanderbilt University
University of Virginia
University of Washington
Washington State University
Washington University in St. Louis
University of Waterloo
Wayne State University
University of Western Ontario
University of Wisconsin–Madison
Yale University
York University