

Census of Institutional Repositories in the United States

MIRACLE Project Research Findings

by Karen Markey, Soo Young Rieh, Beth St. Jean,
Jihyun Kim, and Elizabeth Yakel

February 2007

Council on Library and Information Resources
Washington, D.C.

ISBN 978-1-932326-28-4
CLIR Publication No. 140

Published by:

Council on Library and Information Resources
175 Massachusetts Avenue, NW, Suite 500
Washington, DC 20036

Web site at <http://www.clir.org>

This publication is available online at no charge at <http://www.clir.org/pubs/abstract/pub140abst.html>.

Copyright 2007 by the Council on Library and Information Resources. No part of this publication may be reproduced or transcribed in any form without permission of the publishers. Requests for reproduction or other uses or questions pertaining to permissions should be submitted in writing to the Director of Communications at the Council on Library and Information Resources.

Contents

List of Illustrations	v
About the Authors	vii
Authors' Acknowledgments	viii
Foreword	ix
Abbreviations	xi
Executive Summary	1
1 Background and Methodology	7
1.1 The Impetus for the MIRACLE Project's Census of IRs in the United States	7
1.2 Obtaining a Mailing List of Academic Library Directors	8
1.3 Conducting Comparative Analysis of Survey Software	8
1.4 Drafting and Pretesting Survey Instruments	9
1.5 Setting Up Survey Administration Procedures and Protocol	9
1.6 Data Collection	10
1.7 Data Analysis	12
1.8 Chapter 1 Summary	12
2 The Institutions and the People Involved with IRs	13
2.1 Census Respondents	13
2.2 Using CCHE to Characterize Participating Institutions	16
2.3 Positions of the People Involved with IRs	18
2.4 Number of People Involved with IRs	20
2.5 Positions of People Involved with IRs	20
2.6 The Responsibility for the IR	23
2.7 Involvement with IRs	23
2.8 Chapter 2 Summary	24
3 The Budget for an IR	25
3.1 Sources of Funding	25
3.2 IR Budget Line Items	27
3.3 Chapter 3 Summary	28
4 Important Investigative Activities	29
4.1 Important Investigative Activities	29
4.2 The Needs Assessment	31
4.3 Pilot Testing IR Software Packages	32
4.4 Next Steps Pertaining to the IR Effort	34
4.5 Chapter 4 Summary	37
5 IR Systems and Features	38
5.1 Number of IRs at Institutions	38
5.2 IR Software	38
5.3 IR-system Features	40
5.4 Migrating to a New IR	41
5.5 Chapter 5 Summary	42

6	IR Practices and Policies	43
6.1	The Number of Digital Documents in IRs	43
6.2	Digital Document Types in IRs	44
6.3	Status of IR Policies	46
6.4	File Formats Guaranteed in Perpetuity	50
6.5	Contributors to IRs	52
6.6	Chapter 6 Summary	57
7	IR Benefits and Beneficiaries	58
7.1	Benefits of IRs	58
7.2	Building Relationships	61
7.3	Deploying a Successful IR	62
7.4	IR Contributors	63
7.5	Evaluation Methods	65
7.6	Chapter 7 Summary	66
8	Institutions That Have No Involvement with IRs	67
8.1	Reasons for No Planning	67
8.2	IR-related Activities	69
8.3	Likelihood of IR Planning	72
8.4	Chapter 8 Summary	72
9	Discussion of Census Findings	74
9.1	The Sleeping Beast of Demand for IRs from Master's and Baccalaureate institutions	74
9.2	Verifying Previous Survey Findings Pertaining to Institutions Involved with IRs	78
9.3	Building on Our Knowledge of IRs	79
9.4	Observations on Long-term Issues Pertaining to IRs	85
9.5	Chapter 9 Summary	90
	APPENDIX A. Advisory Group Members	92
	APPENDIX B. Questionnaire for No Planning (NP) Respondents	93
	APPENDIX C. Questionnaire for Planning Only (PO) Respondents	100
	APPENDIX D. Questionnaire for Planning and Pilot Testing (PPT) Respondents	112
	APPENDIX E. Questionnaire for Implementation (IMP) Respondents	129
	APPENDIX F. Literature Review	147
F1	IR Definition and Attributes	147
F2	The First IR Systems, IR Projects, and IR-related Projects	148
F3	IR Surveys and the Extent of Survey Respondents' Involvement with IRs	149
F4	The People Involved with IRs	150
F5	The Budget for an IR	152
F6	Important IR Investigative Activities	153

F7	IR-system Software, Content, and Features.	153
F8	IR Practices and Policies.	154
F9	IR Benefits and Beneficiaries	157
F10	Institutions Where No IR Planning Has Been Done	158
	REFERENCES	160

Illustrations

Tables

1.1	Data collection rounds	11
2.1.	Respondents' positions based on the extent of IR involvement at their institutions	15
2.2	Classes of CCHE institutions invited to participate in the MIRACLE census	17
2.3	CCHE classes reveal the extent of IR involvement by census respondents	18
2.4	Positions of people involved in the IR effort.	19
2.5	Positions of people leading the IR effort	21
3.1	Top- and bottom-ranked funding sources.	26
4.1	Important investigative activities	29
4.2	Importance of the needs assessment	32
4.3	Important benefits of pilot testing	33
4.4	Next steps pertaining to the IR effort.	34
5.1	Number of IRs	38
5.2	Pilot-tested and implemented IRs.	39
5.3	Ranking IR-system features.	40
5.4	Reasons for migrating to a new system	41
6.1	Document types in pilot-test and operational IRs	44
6.2	Guaranteed digital file formats	50
6.3	Authorized contributors to IRs	52
6.4	The major contributor to the IR.	53
7.1	IR benefits.	60
7.2	Increases in benefits' importance	61
7.3	Factors inhibiting the deployment of a successful IR	62
7.4	Reasons for contributing to the IR	64
7.5	Methods of assessing the IR's success	65
8.1	Reasons for no planning.	67
8.2	IR-related activities on the road to an IR	70
8.3	What would have to happen for IR planning to begin at your institution?	71
9.1	How the MIRACLE Project can assist NP respondents regarding IRs	77
9.2	Previous survey findings verified in the MIRACLE Project census	78
9.3	Characterizing respondents' requests for more information about IRs	80

Figures

2.1	Extent of IR involvement	14
2.2	Respondents who completed questionnaires	14
2.3	MIRACLE institutions versus CCHE institutions	17
2.4	Number of people involved in the IR effort	20
2.5	IR committee membership.	21

2.6	Responsibility for the IR.....	23
2.7	Involvement with IRs.....	24
3.1	IR budget line items.....	27
4.1	Did your institution conduct a needs assessment?.....	31
4.2	Likelihood of PO institutions pilot testing.....	34
4.3	Top-ranked next steps.....	35
4.4	Middle-ranked next steps.....	35
4.5	Bottom-ranked next steps.....	36
4.6	Time needed to make the IR implementation decision.....	36
5.1	Likelihood of modifying IR software.....	41
6.1	Number of digital documents in IRs.....	43
6.2	Status of IR policies at IMP institutions.....	47
6.3	Status of IR policies at PPT institutions.....	48
6.4	Managing IP rights in the IR.....	49
6.5	Recruiting content by working with early adopters.....	55
6.6	Recruiting content by word of mouth.....	55
6.7	Recruiting content by personal visits and presentations.....	55
7.1	Increasingly positive nature of respondents' ratings for IR benefits.....	59
7.2	Effect on building relationships.....	62
7.3	Ease of getting people to contribute to the IR.....	63
8.1	Low-ranked reasons for no IR planning.....	68
8.2	Likelihood of IR planning.....	72

About the Authors

Karen Markey is a professor in the School of Information (SI) at the University of Michigan (U-M). Prior to joining Michigan's faculty in 1987, she was a senior research scientist at OCLC. Karen has received research funding from the Council on Library Resources, the U.S. Department of Education, Forest Press, the Institute of Museum and Library Services, the National Science Foundation, and OCLC. The author of four books, more than a dozen major research reports, and dozens of journal articles and conference papers, she has been invited to speak at meetings in North America, Europe, and Australia. She is a coprincipal investigator of the MIRACLE Project.

Soo Young Rieh is an assistant professor at U-M SI. She teaches in the areas of information-seeking behavior, use of information, and evaluation of systems and services. Soo's research seeks to better understand people's interactions with information in various contexts, such as the Web, libraries, and institutional repositories. Her primary research interests lie in credibility- and cognitive-authority judgments in the information-seeking process. With regard to the MIRACLE Project, Soo is particularly interested in identifying success factors in institutional repositories from multiple perspectives, including administration, service, and use. She is principal investigator of the MIRACLE Project.

Beth St. Jean is a first-year doctoral student at U-M SI. She holds a bachelor's degree in mathematics from Smith College and a master's degree in information with a specialty in library and information services from U-M SI. Before coming to U-M, Beth worked in the fields of financial and statistical analysis for more than 15 years. She is a graduate student research assistant on the MIRACLE Project.

Jihyun Kim is a doctoral candidate at U-M SI. She has been involved in research projects on access systems for archival information, particularly research funded by National Historical Publications and Records Commission on the usability of archival access tools. Jihyun is writing her dissertation, which investigates factors that motivate and impede faculty self-archiving behavior. She is a research assistant on the MIRACLE Project.

Elizabeth Yakel is an associate professor at U-M SI, where she teaches in the Archives and Records Management Specialization. Her research interests include analyzing archival user needs and improving access to primary sources. Beth has received grants from The Andrew W. Mellon Foundation, the National Historical Publications and Records Commission, and OCLC. She has published on many aspects of archival use and user services in such publications as *American Archivist*, *Archivaria*, and *Archival Science*. She is a coprincipal investigator on the MIRACLE Project.

Authors' Acknowledgments

In this report, we describe results of a nationwide census of institutional repositories in U.S. academic institutions. The census is one of several activities of the MIRACLE (Making Institutional Repositories a Collaborative Learning Environment) Project. The Institute of Museum and Library Services (IMLS) provides support for the MIRACLE Project through its National Leadership Grants Program (grant number LG01-05-0126). Martha Crawley is program officer. Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily represent those of the IMLS.

The MIRACLE Project staff thanks members of the advisory committee—Joseph Branin, Michael Seadle, Helen Tibbo, Diane Vizine-Goetz, and Marcia Zeng—who visited Ann Arbor or provided telephone advice on the construction of the census, the content of this report, and project business in general. We are also grateful to Yong-mi Kim, who took part in our early discussions about census questions, participants, and Web survey software.

We thank Susan L. Perry at CLIR for responding positively to our inquiries to publish MIRACLE Project census findings in CLIR's publication series. Now these findings will reside at an electronic location where prospective readers are likely to find additional information on related and allied topics. At CLIR, Kathlin Smith gave us superb editorial support, streamlined production of the report, and coordinated her efforts with Brian Leney, who performed design and layout tasks.

We are especially grateful to the many staff at educational institutions throughout the country who took part in the census.

Foreword

The subject of institutional repositories commands great interest on campuses across the country, and for good reason. At the heart of higher education is the generation and dissemination of knowledge. It is only natural that campus leaders, witnessing the startling proliferation of new information made possible by digital technologies, are growing concerned about the stewardship of the knowledge assets produced in their institutions. For many academic leaders, institutional repositories seem an ideal tool to manage knowledge production and dissemination. There is a growing need for carefully gathered evidence that will help people learn about the growth and effectiveness of repositories.

This report is designed to meet that need. It presents and analyzes data gathered in the first phase of the MIRACLE Project, an IMLS-funded research program based at the University of Michigan that is “investigating the implementation of institutional repositories in colleges and universities in order to identify models and best practices in the administration, technical infrastructure, and access to repository collections.” In the first phase, MIRACLE Project investigators conducted a nationwide census of institutions to determine the extent of their involvement with repositories. This data-gathering phase will be followed by a series of telephone interviews with institutional repositories’ staff, in-depth investigations into five institutional repositories, a survey of repository users, and, finally, a study based on searching in repositories.

In part because the census revealed a great demand for information about what is going on in the world of institutional repositories, the MIRACLE researchers decided to publish their initial data before subsequent phases of the project are done. Like all satisfying surveys, this one not only confirms what we already know but also introduces uncertainties where once there was seeming clarity.

A conspicuous fact about institutional repositories, confirmed by the MIRACLE Project findings, is that there is no consensus on what institutional repositories are for. Many librarians and administrators are convinced that repositories are important—so much so that most are, or will be, implementing repositories before they do a needs assessment. The investigators found that when such an assessment is done, it often occurs *after* an organization has decided to implement a repository and that its goal is to predict adoption rates by targeted users.

While motivations for implementing repositories vary, some expectations appear to be at odds with the results reported by early adopters. For example, many institutions that plan or pilot test repositories are motivated by the desire to change the dynamics of scholarly communication. Yet the survey confirms a finding, reported elsewhere anecdotally, that operating repositories have had limited success in recruiting voluntary deposit of content. Other institutions identify stewardship of digital assets, especially their preservation, as a key function of a repository. Yet survey data confirm that repositories are not yet providing key preservation services, such as guaranteeing the integrity of file formats for future use. It is one of the paradoxical findings of the survey that there is a detectable urgency on the part of libraries to implement institutional repositories, even as early adopters report difficulties in achieving the purposes for which they built them.

These results raise several interesting questions that subsequent phases of this project are likely to illuminate. Why haven’t more faculty and researchers contributed to institutional repositories? Are institutional repositories what scholars need, or will most of them choose to deposit their materials in domain-

specific repositories, if and when such facilities are developed? Is the ultimate purpose of an institutional repository to ensure strong institutional stewardship of records over time, rather than to facilitate current access?

The idea of stewardship itself merits additional investigation. Why do many see institutional repositories as part of the solution to the challenges of digital preservation, even though none of the repositories now in place reports that they provide reliable format-specific preservation (with perhaps the exception of preserving PDFs)? Five years hence, it would be interesting to see whether those institutions that now say they will change their software to address the preservation issue have actually been able to do so. Another question is why so few archivists are involved in the design, implementation, and management of repositories given that institutions are targeting unpublished materials for deposit.

The next few years will be critical in the fate of institutional repositories. Proponents see them developing into a key element of the new technological infrastructure that campuses require. The retooling of existing infrastructure into what is widely called the “cyberinfrastructure” requires that all components of the existing infrastructure—including libraries (as publishers) and archives—be rethought, retooled, and repurposed. In this context, it may not be surprising that there is a gap between the claims of stewardship—or aspirations for stewardship—by institutional repositories and their current ability to preserve digital assets. Organizational models for digital preservation are only now emerging, and they are quite diverse. They range from all-purpose storage solutions for high-use, high-volume data (such as some programs in the San Diego Supercomputer Center) to highly selective, closed systems that keep proprietary assets dark (such as Portico). Several scenarios are possible. In some disciplines, institutional repositories may play significant roles in disseminating both unpublished and published research results. In other disciplines, institutional repositories could have no significance whatsoever; scholars in these fields will continue to prefer to deposit their materials in domain-based archives. The differences will arise not from institutional arrangements but from the nature of the content and of the communities who use it.

CLIR’s attention to institutional repositories is an extension of its long-standing interest in libraries’ support of research and learning. As new models of repositories emerge and are tested, all stakeholders will need ready access to information about them. It will also be important that these repositories can interoperate or, at a minimum, that they expose their content’s metadata. Regardless of the organizational models that ultimately hold, the policies that govern them will be of paramount importance.

It may be true, as this report indicates, that people are enthusiastic about institutional repositories because they see them as a technology-based solution to a number of challenges in the digital environment. But people, not technology, make solutions. And to find solutions, it is best to start with the facts. It is in that spirit that CLIR offers this report.

Abby Smith

Abby Smith currently works with the Library of Congress’s National Digital Information and Preservation Program (NDIIPP) in development of its national strategy to identify, collect, and preserve digital content of long-term value. She was director of programs at CLIR until she relocated to California in 2005.

Abbreviations**ALD**

American Library Directory

ARL

Association of Research Libraries

CARL

Canadian Association of Research Libraries

CCHE

Carnegie Classification of Institutions of Higher Education

CIO

chief information officer

CNI

Coalition for Networked Information

CV

curriculum vitae

DARPA

Defense Advanced Research Projects Agency

ETDs

electronic theses and dissertations

FAIR Programme

Focus on Access to Institutional Resources Programme

FEDORA

Flexible and Extensible Digital Object and Repository Architecture

FTE or FTEs

full-time equivalent(s)

GNU

GNUs not UNIX

ICPSR

Inter-University Consortium for Political and Social Research

IHE

institution of higher education

IMLS

Institute of Museum and Library Services

IMP

implementation (respondents in the MIRACLE Project census)

IP

intellectual property

IR

institutional repository

JISC

Joint Information Systems Committee

KM

knowledge management

LEADIRS

Learning About Digital Institutional Repositories

LITA

Library and Information Technology Association

METS

Metadata Encoding and Transmission Standard

MIRACLE Project

Making Institutional Repositories a Collaborative Learning Environment Project

MIT

Massachusetts Institute of Technology

NSF

National Science Foundation

NP

no planning (respondents in the MIRACLE Project census)

OA

open access

OAI

Open Archives Initiative

OAI-PMH

Open Archives Initiative Protocol for Metadata Harvesting

OCLC

OCLC Online Computer Library Center, Inc.

OAIS Reference Model

Open Archival Information System Reference Model

OSU

The Ohio State University

PDF

portable document file

PPT

planning and pilot testing (respondents in the MIRACLE Project census)

PO

planning only (respondents in the MIRACLE Project census)

RLG

Research Libraries Group

ROARMAP

Registry of Open Access Repository Material Archiving Policies

SI

School of Information (University of Michigan)

SPEC Kit

System and Procedures Exchange Center Kit

SPSS

Statistical Package for the Social Sciences

TDR

trusted digital repository

UK

United Kingdom

U-M

University of Michigan

URL

Uniform Resource Locator

XML

extensible markup language

EXECUTIVE SUMMARY

Why Study Institutional Repositories?

A considerable portion of the scholarly record is born digital, and some scholarship is produced in digital formats that have no physical, in-the-hand counterparts. The proliferation of digital scholarship raises serious and pressing issues about how to organize, access, and preserve it in perpetuity. The response of academic institutions has been to build and deploy institutional repositories (IRs) to manage the digital scholarship their learning communities produce. IR efforts require a considerable financial, personnel, and technical investment. For this reason, it would be helpful if academic institutions could learn from one another, sharing their experiences, building models, and formulating best practices. Such sharing would streamline the implementation of IRs at institutions where the decision to initiate an IR effort has not yet been made.

Why Conduct a Census of IRs in the United States?

Previous surveys have focused on academic institutions where IRs are already operational or on specialized groups of academic institutions that are likely to be first adopters of new technologies such as IRs (Appendix F3). To avoid duplication, MIRACLE Project staff (i.e., this report's authors) sought to cast a wide net and fill a void. Conducting a *census* of academic institutions in the United States about their involvement with IRs would include institutions that have not yet jumped on the IR bandwagon. Being inclusive increases our confidence that we will be able to identify the wide range of practices, policies, and operations in effect at institutions where decision makers are contemplating, planning, pilot testing, or implementing IRs and will enable us to learn why some institutions have ruled out IRs entirely.

Who Participated in the MIRACLE Project Census of IRs in the United States?

Of the 2,147 academic library directors and senior library administrators MIRACLE Project staff contacted, 446 participated in the census—a response rate of 20.8%. Characterizing the extent of their involvement with IRs, 236 (52.9%) respondents reported that they have done no IR planning (NP) to date, 92 (20.6%) respondents are only planning (PO) for IRs, 70 (15.7%) respondents are actively planning and pilot testing IRs (PPT), and 48 (10.8%) respondents have implemented (IMP) an operational IR (Figure 2.1).

What Kinds of Educational Institutions Have and Do Not Have IRs?

MIRACLE Project staff used the Carnegie Classification of Institutions of Higher Education (CCHE) to characterize census respondents (Table 2.2 and Figure 2.3). Research universities vastly outnumber other CCHE classes with respect to involvement in IR planning, pilot testing, and implementation (Table 2.3). Most NP and PO respondents come from master's and baccalaureate institutions.

Who Bears the Responsibility for IR Planning, Pilot Testing, and Implementation?

At PPT and IMP institutions, librarians take the lead in IR pilot testing and system implementation (Table 2.4), assume most of the responsibility for the IR effort (Figure 2.6), and are members of various IR committees (Figure 2.5). Funding almost always comes from the library (Table 3.1). A typical approach to funding the IR is to absorb its cost in routine library operating costs.

At NP institutions where no IR effort is under way, the library director takes the lead, consulting with the provost, chief information officer, faculty, and archivist about funding, technical expertise, potential contributors and users, and digital collections (Tables 2.4 and 2.5). IR committee membership becomes increasingly less inclusive as the IR project progresses from pilot testing to implementation, leaving the library "holding the bag" (Figure 2.5).

What Are Useful Investigative Activities?

Staff involved with various phases of IR efforts have voracious appetites for information about IRs, especially information pertaining to best practices and successful implementations at institutions similar to their own (Tables 4.1, 8.1, 8.2, and 9.3). The needs assessment is not as important as other investigative activities (Table 4.1 and Figure 4.1). Pilot testing one or more IR-system packages is very important. About 16% of MIRACLE census respondents are pilot testing one or more IR-system packages (Figure 2.1), and almost three-quarters of PO respondents intend to pilot test IR-system software (Figure 4.2). Benefits of pilot testing include developing the requisite technical expertise for IR implementation, evaluating IR-system software, and estimating implementation costs (Table 4.3). For most PO institutions in the census, the next step is to widen the scope of their investigations. For most PPT institutions, the next step is to implement IR-system software (Figure 4.3). Very few (about 10%) PO and PPT institutions are likely to terminate their IR efforts (Figure 4.5).

What Are Respondents' Experiences with IR-system Software Packages?

Respondents' preferred IR-system software for both pilot testing and implementation is DSpace (Table 5.2). Asked how long their IR has been operational, 52.1% of respondents with operational IRs cite

12 months or less, 27.1% from 13 to 24 months, 4.2% from 25 to 36 months, and 16.6% for more than 36 months. IR-system functionality is satisfactory, but the user interface, including controlled vocabulary searching and authority control, needs serious reworking (Table 5.3). Except for portable document files (PDFs), institutions with operational IRs do not guarantee file formats in perpetuity (Table 6.2). Improving preservation functionality in IRs should be a systems-development priority because IMP respondents rate greater preservation capacity as *the* major reason why they will migrate to a new IR (Table 5.4). To date, respondents have used IR-system evaluation methods that are limited to simple counts that most IR systems produce automatically in management reports (Table 7.5).

What Content Is in Pilot-test and Operational IRs?

Both pilot-test and operational IRs are very small (Figure 6.1). About 80% of the former and 50% of the latter contain fewer than 1,000 digital documents. Only four (8.3%) pilot-test IRs and seven (19.4%) operational IRs contain more than 5,000 documents. There is no relationship between IR size and age. Pilot-test and operational IRs contain a wide range of text, numeric, and multimedia files, but traditional text-based document types that are the result of the research enterprise of staff and students at postsecondary institutions are especially characteristic of these institutions' content (Table 6.1).

What Progress Have Respondents Made on IR Policies?

At least 60% of census respondents with operational IRs report they have implemented policies for (1) acceptable file formats, (2) determining who is authorized to make contributions to the IR, (3) defining collections, (4) restricting access to IR content, (5) identifying metadata formats and authorized metadata creators, and (6) determining what is acceptable content (Figure 6.2). There are many more IR-related activities for which these institutions report drafted policies or no policies at all.

It may be not necessary for all IR policies to be in place at the time of the public launch of an institution's IR. Taking a wait-and-see attitude, evaluating what transpires after a period of time, then firming up existing policies and implementing new ones as needed may be the most expedient course of action.

Who Contributes to IRs and at What Rate?

Authorized contributors to IRs are typically members of the institution's learning community—faculty, librarians, research scientists, archivists, and graduate and undergraduate students (Table 6.3). Staff who facilitate the research and teaching missions of the institution (e.g., press, news service, academic support staff, central computer staff) are less likely to be authorized to contribute. Asked to identify *the* major contributor to their IR, only PPT staff are uni-

fied in their response, with almost 60% naming faculty (Table 6.4). Percentages drop to 48.1% and 33.3% for PO and IMP respondents, respectively. The unified response of PPT staff probably stems from the fact that they work one-on-one with faculty who are early adopters during the planning and pilot-test phase of the IR effort. In fact, PO, PPT, and IMP respondents choose “IR staff working one-on-one with early adopters” as the most successful method for recruiting IR content (Figure 6.5). Other successful methods are “word of mouth from early adopters to their colleagues” (Figure 6.6), “personal visits to staff and administrators,” and “presentations about the IR at departmental and faculty meetings” (Figure 6.7).

Respondents report that recruiting content for the IR is difficult (Figure 7.3). At institutions with operational IRs, IR staff are willing to entertain institutional mandates that require members of their institution’s learning community to deposit certain document types in the IR (Table 7.3). Asked why they think people will contribute to the IR, respondents give high ratings to reasons that enhance scholarly reputations and offload research-dissemination tasks onto others. Lower-ranked reasons pertain to enhancing the institution’s standing.

What Are the Benefits of IRs?

Asked to rate a list of 14 benefits of IRs, census respondents give high ratings to all but two (Figure 7.1 and Table 7.1). Instead of having a few benefits that stand far above the others, IRs may have many benefits. Respondents may also feel it is premature to rank one or two benefits above the others because IRs have not yet “come into their own.” Once IRs have become more common in all types of educational institutions, the answers to this question might be different. One or two benefits may ultimately dominate.

NP respondents are especially interested in benefits of IRs so they can incorporate them into arguments to convince their institutions’ decision makers to support IR planning (Tables 8.2 and 9.1).

What Factors Inhibit the Deployment of a Successful IR?

Factors affecting the successful deployment depend on the stage of an institution’s IR effort (Table 7.3). IMP respondents are concerned about contributors and contributions to the IR. In fact, that concern is pushing them to consider mandating contributions of certain material types. PPT respondents are also concerned about contributions, but other priorities, projects, and initiatives are competing with the IR effort for resources. PO respondents are most concerned about sustaining the IR effort in terms of competing for resources and supporting the costs of an operational IR.

How Likely Are Institutions Where No IR Planning Has Been Done to Jump on the IR Bandwagon?

The largest percentage (52%) of MIRACLE Project census respondents comes from institutions where no IR planning has been done. Dominating these NP respondents are master's and baccalaureate institutions (Table 2.3).

Among NP respondents is a sleeping beast of demand for IRs. These respondents want to know how much IRs cost to plan, implement, and maintain, and what institutions comparable to their own are doing with regard to IRs (Table 8.2 and Subchapter 9.1). None of the top-ranked reasons why NP institutions have not begun IR planning rules out future involvement with IRs (Table 8.1); however, right now, NP institutions have other things on their plate or have insufficient resources or expertise for IR planning. Very few are totally in the dark in terms of what IRs are and whether IRs have relevance for their institutions (Figure 8.1). Slightly under 50% of NP respondents may start IR planning within the next 24 months (Figure 8.2).

Asked how the MIRACLE Project could assist them regarding IRs, NP respondents want to learn about (1) IRs generally, (2) the details and specifics of IRs, (3) best practices, (4) benefits of IRs, (5) securing funding for IRs, and (6) opportunities for partnerships (Table 9.1). NP respondents' interest in IRs is a wake-up call to their colleagues at other-than-research-universities to share their success stories about IRs with an audience that is craving for information. It is also an opportunity for the MIRACLE Project to focus on other-than-research-universities in subsequent project activities because that is where the need is greatest and where the gap in our knowledge about IRs is widest.

What Previous Findings about IRs Do MIRACLE Project Census Findings Verify?

The MIRACLE Project census verifies almost two dozen findings from previous surveys. Among these findings are that research universities lead in the implementation of IRs, that libraries play a leading role in the IR effort, and that DSpace leads in IR-system pilot testing and implementation. See Table 9.2 for the complete list.

What Findings Are Unique to the MIRACLE Project Census?

Subchapter 9.3 features a discussion of 13 unique findings. Examples are the shrinking-violet role that archivists play in the IR effort; the voracious appetites that census respondents have for information especially about successful IR implementations at institutions similar to their own; the ability of the IR to forge new relationships for libraries; and the need for improved preservation functionality in IRs.

What Long-term Issues Will Occupy IR Staff Long after the MIRACLE Project Ends?

Subchapter 9.4 discusses seven such issues. Examples are the benefits of IRs, the effect of IRs on derailing the current publishing model, and requiring learning communities to submit the products and by-products of their research and teaching enterprises to the IR.

1 BACKGROUND AND METHODOLOGY

Chapter 1 gives background on the MIRACLE Project, defines institutional repositories (IRs), and describes the methods MIRACLE Project staff used to conduct a census of IRs in U.S. academic institutions.

1.1 The Impetus for the MIRACLE Project's Census of IRs in the United States

A considerable portion of the scholarly record is born digital, and some scholarship is produced in digital formats that have no physical, in-the-hand counterparts. The proliferation of digital scholarship raises serious and pressing issues about how to organize, access, and preserve it in perpetuity. The response of U.S. colleges and universities has been to build IRs to capture, preserve, and reuse the intellectual output of teaching, research, and service activities at their institutions. An IR is “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members” (Lynch 2003) (see also Appendix F1).

The MIRACLE (Making Institutional Repositories a Collaborative Learning Environment) Project is investigating the implementation of IRs at academic institutions to identify models and best practices for the administration, technical infrastructure, and access to digital collections. The chief objective of the project is to identify specific factors contributing to the success of IRs and effective ways of accessing and using IRs. The census is the first of several activities aimed at achieving project objectives. Other activities will study IR users, contributors, and staff through the use of telephone interviews, case studies, personal interviews, observations, and experiments.

Originally, MIRACLE Project investigators proposed to survey *operational* IRs in North America; however, we were concerned that we would be duplicating the efforts of Charles Bailey and his University of Houston associates who were analyzing data from their Association of Research Libraries (ARL)-sponsored survey of member institutions at the same time we were making data-collection decisions for the MIRACLE survey (Bailey et al. 2006). Other surveys targeted specific user groups such as Coalition for Networked Information (CNI) members in the United States (Lynch and Lippincott 2005), CNI members abroad (van Westrienen and Lynch 2005), Canadian Association of Research Libraries (CARL)-member libraries (Shearer 2004), and early adopters of IR technology worldwide (Mark Ware Consulting 2004).

Examining these surveys' results, MIRACLE project investigators decided not to limit their efforts to a particular user group, membership, or affiliation, and not to restrict participation to institutions with an operational IR. Instead, we sought to cast our net broadly and fill a void. Conducting a *census* of academic institutions in the United States about their involvement with IRs, MIRACLE Project investigators decided not to exclude institutions that have not jumped on the IR bandwagon. Being more inclusive would increase our confidence that we would be able to identify the wide range of practices, policies, and operations in effect at institutions where decision makers are contemplating, planning, pilot testing, or implementing IRs. At the same time, it would enable us to learn why some institutions have ruled out IRs entirely.

1.2 Obtaining a Mailing List of Academic Library Directors

The first task of MIRACLE Project staff was to obtain an electronic mailing list bearing the names and e-mail addresses of academic library directors and senior library administrators at U.S. educational institutions. A number of companies provide this information for a fee (for example, see American Library Association 2006). After examining their products and services, MIRACLE Project staff narrowed options to the following four companies or products: (1) Thomson-Peterson's, (2) Market Data Retrieval, (3) American Library Directory Online, and (4) World Guide to Libraries Plus. After comparing these companies' products with respect to such variables as the number of records with e-mail addresses available, scope, and price, as well as other advantages and disadvantages, we decided to purchase mailing lists from two vendors: (1) American Library Directory (ALD) and (2) Thomson-Peterson's. Using ALD's online database, we downloaded a comprehensive list (2,207 records) of all college and university main libraries in the United States (including U.S. protectorates). Because ALD's online database did not provide e-mail addresses for specific individuals, we purchased a less comprehensive database from Thompson-Peterson's that we used to add e-mail addresses to ALD data. After deleting community colleges and duplicates, we ended up with 2,147 e-mail addresses for the nationwide census.

1.3 Conducting Comparative Analysis of Survey Software

To compare survey-software programs for administering our Web-based survey, MIRACLE Project staff signed up for free trials of 10 such software programs: SurveyMonkey, Zoomerang, Key Survey, SurveyConsole, EZQuestionnaire, iSalient, QuestionPro, Ridgcrest Surveys, SmartSurveys, and SuperSurvey. Staff also researched Flashlight Online, ScyWeb, and UM.Lessons. On the basis of pricing information, flexibility, and functionality, we narrowed the list

to SurveyMonkey, Zoomerang, Key Survey, and UM.Lessons. Staff eliminated UM.Lessons and Key Survey from consideration because of the former's limited flexibility and functionality and the latter's cost.

MIRACLE Project staff's decision to use SurveyMonkey instead of Zoomerang was based on the former program's greater flexibility and functionality. Our purchase of a one-year professional subscription to SurveyMonkey would enable us to launch an unlimited number of surveys with an unlimited number of questions and to use its advanced features for the survey's many complicated questions.

1.4 Drafting and Pretesting Survey Instruments

To draft survey instruments, MIRACLE Project investigators reviewed published and open-access literature on IRs through 2005 (see the MIRACLE Project's bibliography for a list of relevant publications at <http://miracle.si.umich.edu/bibliography.html>), talked to colleagues, and asked advisory group members (see Appendix A) to review, comment on, and edit draft instruments. Because the investigators expected survey respondents to come from institutions that were at various stages of the IR effort, they could neither ask everyone the same questions nor ask questions in the same way. Advice from advisory group members resulted in these four categories of IR involvement: (1) no planning to date (NP), (2) planning only to date (PO), (3) planning and pilot testing one or more IR systems (PPT), and (4) public implementation of an IR system at the respondent's institution (IMP). MIRACLE Project investigators drafted four different questionnaires based on these four categories of IR involvement.

Asking the same or similar questions in two or more questionnaires would enable investigators to make comparisons among institutions on the basis of the extent of their involvement with IRs. For example, here is a question about anticipated benefits of IRs that is worded a little differently depending on an institution's involvement with IRs:

- For NP respondents: How important do you think these anticipated benefits of IR would be to your institution?
- For PPT and PO respondents: How important are these anticipated benefits of IR to your institution?
- For IMP respondents: At the beginning of IR planning at your institution, how important did you think these anticipated benefits of IR would be to your institution?

Appendixes B, C, D, and E contain the MIRACLE Project questionnaires for NPs, POs, PPTs, and IMPs, respectively.

1.5 Setting Up Survey Administration Procedures and Protocol

Having so many institutions (2,147) in the census sample required MIRACLE Project staff to work out a detailed distribution plan. After pretesting a few different approaches, we decided to send an

e-mail message to each institution's academic library director or a senior administrator to tell them about the census and to ask them about the extent of their involvement with IRs. More specifically, we wrote, "Please tell me how you would characterize the current status of your institutional repository (IR)." We asked them to base their response on one of four categories: (1) no planning to date, (2) planning only to date, (3) both planning and pilot testing one or more IR systems, and (4) public implementation of an IR system at their institution.

On the basis of the person's response, we replied with an e-mail message bearing a link to the appropriate Web-administered questionnaire (see Appendixes B, C, D, and E for NP, PO, PPT, and IMP questionnaires, respectively). We used SurveyMonkey's list-management tool to send out initial survey links and to perform two subsequent follow-ups with individuals who had agreed to participate but who had failed to respond to our inquiries.

Recruiting people to participate in the MIRACLE census in this way is the electronic version of what those in the sales world term a "cold call." We sent prospective respondents e-mail messages with a substantive phrase in the "SUBJECT" line announcing our IR census and asked them to participate. It is likely that the people who responded to our e-mail message were interested in IRs and thus were more likely to open, read, and respond to such a message and eventually respond positively about IRs on their completed questionnaires. Thus, MIRACLE census respondents may be more favorably inclined toward IRs than other academic library directors and senior administrators generally because of how we recruited them.

1.6 Data Collection

MIRACLE Project staff conducted the nationwide IR census from April 19, 2006, through June 24, 2006. Data collection was not straightforward. When few respondents responded to our invitations and reminders, we discussed problems and brainstormed ways of solving them. For example, coprincipal investigator Elizabeth Yakel suggested replacing the original SUBJECT line in our e-mail messages, "IMLS Institutional Repositories Census," with the catchier phrase, "Be Counted! National Institutional Repository Census." This change did indeed result in a higher response rate.

Table 1.1 summarizes the six data collection rounds that were necessary to increase the survey's invitational response rate to an acceptable level.

Concurrent with sending e-mail invitations, MIRACLE Project staff e-mailed a link to the appropriate Web-administered questionnaire to respondents within three business days of their response to our invitation. When respondents failed to return the completed questionnaires, staff sent them up to two reminders. The text of these two e-mail responses (the first survey link e-mail and the reminder e-mail) remained fairly stable throughout the census. Staff took care to

Table 1.1 Data collection rounds

No. of invitations sent	Date	Cumulative invitation response rate*		Cumulative survey response rate†		Explanations and changes made to increase response rates
		No.	%	No.	%	
2,147	4/19 to 4/26	172	9	89	5	Invitations sent through Rieh's e-mail account. Staff research 260 undeliverable messages.
1,698	5/2 to 5/14	320	15	169	8	Invitations sent through Markey's e-mail account. Staff continue to research undeliverable messages.
1,805	5/15 to 5/22	467	22	273	13	Invitations sent through Markey's e-mail account. Staff change SUBJECT line and invitation text.
1,619	5/23 to 5/30	566	27	370	18	Invitations sent through Markey's e-mail account.
1,511	5/31 to 6/7	627	30	420	20	Invitations sent through Yakel's e-mail account.
1,446	6/8 to 6/24	676	32	500	24	Yakel's account. Staff change SUBJECT line announcing end of census. Seven undeliverable messages.

* Total number of people who responded to our invitation stating that they were willing to participate in the MIRACLE Project census.

† Total number of people who clicked on the SurveyMonkey link that MIRACLE Project staff sent to them in response to our invitation. Generally, these figures indicate how many people actually participated in the survey. Because some people who clicked on the link exited the survey without answering any questions, these percentages are inflated. After MIRACLE Project staff had removed empty and nearly empty response sets, deleted duplicates, etc., the census response rate was 20.8%.

send e-mail correspondence from the same account (Rieh, Markey, or Yakel), matching the account to which each respondent had initially responded.

A large number of people who had agreed to participate in the census failed to follow through. To rectify this situation, MIRACLE Project staff drafted two e-mail messages—one for respondents who had not yet started filling out the questionnaire and a second for respondents who had answered some questions. The SUBJECT line of both e-mail messages was "Survey to Close 6/24 (Nationwide Census of Institutional Repositories)." In mid-June, staff sent these e-mails to selected respondents. Because these e-mail messages encouraged a number of respondents to complete questionnaires, staff sent a second message to those who had still not responded and changed the SUBJECT line to "5 Days Left: Last Chance to be Counted in Nationwide Census of Institutional Repositories." Quite a few people filled out questionnaires after receiving the second message. When MIRACLE Project staff closed questionnaire administration in SurveyMonkey at 8 a.m. on June 25, 2006, the invitation response rate was 32%.

1.7 Data Analysis

After closing the census in SurveyMonkey, MIRACLE Project staff exported census data from SurveyMonkey into four Microsoft Excel files (one for each version of the survey—NP, PO, PPT, and IMP). Staff cleaned up census data, deleting the responses of people who did not sign the informed consent form as well as those of people who responded *only* to the informed consent form or to the one question about the number of IRs at their institution. Staff deleted empty questionnaires. They deleted multiple answer sets, keeping only the most comprehensive response sets from respondents. Staff deleted one entry that was submitted from a two-year college. This college had been sent an invitation because of an error in one of the mailing lists that we had purchased. After data cleanup had been completed, the census response rate was 20.8%.

MIRACLE Project staff imported the cleaned-up census data into SPSS and calculated frequency tables for the responses to each question in each of the four survey versions. Using these SPSS calculations, staff created an Excel spreadsheet that depicted frequency tables side-by-side for each question across the four questionnaire versions. Staff also produced a Word document that shows respondents' answers to open-ended questions.

MIRACLE Project staff used related data files to probe research questions in greater depth. For example, they downloaded a file from the Carnegie Foundation's Web site that allowed them to determine whether census participants were representative of educational institutions in the United States (see Subchapter 2.2) (Carnegie Foundation 2006b).

1.8 Chapter 1 Summary

Institutional repositories are the response of U.S. colleges and universities to the problem of organizing, providing access to, and preserving scholarship that their learning communities produce in digital formats.

Originally, MIRACLE Project investigators proposed to survey *operational* IRs in North America; however, we were concerned that we would be duplicating previous surveys that targeted institutions with operational IRs. We decided to cast our net broadly and to conduct a *census* of American academic institutions about their involvement with IRs. Census results would fill a void—yielding data and analyses about educational institutions that are and are not involved with IRs.

MIRACLE Project staff purchased mailing lists from two vendors: (1) ALD, and (2) Thomson-Peterson's. After deleting community colleges and duplicates, we ended with a total of 2,147 e-mail addresses for the nationwide census.

Staff pilot-tested several Web-administered software programs and chose SurveyMonkey because of its flexibility and functionality for the complex questions in MIRACLE questionnaires.

Project investigators drafted questionnaires and received feed-

back from advisory group members regarding questions and response categories. On the basis of their input, staff developed four separate questionnaires based on respondents' extent of involvement with IRs: (1) no planning (NP), (2) planning only (PO), (3) planning and pilot testing (PPT), and (4) implementation (IMP). (See Appendixes B, C, D, and E for NP, PO, PPT, and IMP questionnaires, respectively.)

Data collection took place from April 19, 2006, through June 24, 2006. MIRACLE Project staff sent invitations to participate in the census via e-mail to each institution's academic library director or a senior administrator. The e-mail explained the census and asked them about the extent of their involvement with IRs. We replied via e-mail to those who responded to our request with a link to the appropriate Web-administered questionnaire.

Low response rates to our invitation resulted in changes in the text of our reminder messages, especially the wording of the message's SUBJECT line. After data collection ended on June 24, 2006, MIRACLE Project staff cleaned up census data, for example, deleting empty questionnaires or responses of people who did not sign the informed consent form. After data cleanup had been done, the census response rate was 20.8%. MIRACLE Project staff then proceeded with data analysis and reporting activities.

2 THE INSTITUTIONS AND THE PEOPLE INVOLVED WITH IRs

Chapter 2 examines the extent to which certain types of academic institutions are involved with institutional repositories (IRs) and the nature of people's involvement with IRs at these institutions.

2.1 Census Respondents

Of the 2,147 academic library directors and senior library administrators MIRACLE Project staff contacted, 446 participated in the census—a response rate of 20.8%. Characterizing the extent of their involvement with IRs, 236 (52.9%) respondents have done no IR planning (NP) to date, 92 (20.6%) respondents are only planning (PO) for IRs, 70 (15.7%) respondents are actively planning and pilot testing IRs (PPT), and 48 (10.8%) respondents have implemented (IMP) an operational IR. Figure 2.1 is a graphical representation of the extent of IR involvement by MIRACLE Project census respondents.

When MIRACLE Project staff contacted library directors and senior library administrators by e-mail, we asked them to pass our questionnaire to staff who were most familiar with their institution's involvement with IRs. The questionnaires concluded by asking respondents to identify their positions at their institution. Figure 2.2 shows the titles of those who completed questionnaires.

Fig. 2.1. Extent of IR involvement

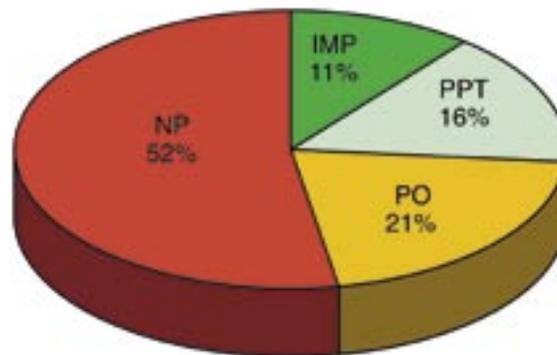
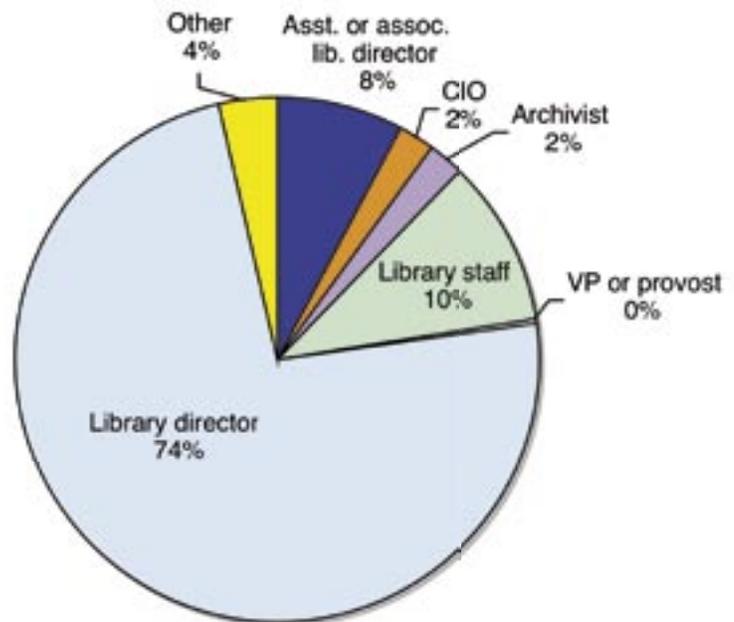


Fig. 2.2. Respondents who completed questionnaires



Almost three-quarters of respondents are library directors; the second- and third-largest percentages (10.2% and 7.9%, respectively) are library staff and assistant-associate librarians, respectively. Library directors prevail in terms of responding to the MIRACLE Project staff's request to participate in the census. We deliberately chose to make library directors or senior library administrators the initial contact at academic institutions because of the difficulty identifying the names of the key person(s) involved with IRs at academic institutions and finding address lists to simplify and streamline contacting tasks. For example, we could have contacted chief information officers (CIOs) instead of librarians but academic institutions do not necessarily apply the CIO moniker across the board nor do all institutions have such a position. The same thing probably applies to ar-

chivists. Even more complicated would have been contacting middle management in academic institutions—deans, directors, chairs, and heads of academic units, research centers, and institutes. Because every academic institution is likely to employ a librarian, we contacted librarians in top management positions to participate in our census.

Our decision to contact librarians may have caused us to miss academic, research, and service units that have implemented or are planning to implement an IR. To some extent, respondents' answers to a census question about how many IRs are available at their institutions may shed light on what we missed (see Chapter 5.1 for answers to this question). MIRACLE Project investigators readily admit that census results may be biased toward libraries because our initial contact was the college or university librarian.

Table 2.1 shows a breakdown of census respondents based on the extent of their institutions' involvement with IRs. At NP institutions, about 90% of respondents are library directors. Percentages in other named-position categories are very small. Of the four people classed in "Other," three are combined library directors-CIOs, and one is head of digital library programs.

Table 2.1. Respondents' positions based on the extent of IR involvement at their institutions

Respondent position	NP		PO		PPT		IMP		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
Library director	194	90.6	57	71.3	29	48.3	8	21.6	288	73.7
Library staff	5	2.3	11	13.8	8	13.3	16	43.3	40	10.2
Assistant or associate library director	5	2.3	0	0.0	16	26.7	10	27.0	31	7.9
Archivist	4	1.9	3	3.7	2	3.3	0	0.0	9	2.3
CIO	1	0.5	5	6.2	1	1.7	1	2.7	8	2.0
VP or provost	1	0.5	0	0.0	0	0.0	0	0.0	1	0.3
Other	4	1.9	4	5.0	4	6.7	2	5.4	14	3.6
Total	214	100.0	80	100.0	80	100.0	37	100.0	391	100.0

At PO institutions, the percentage of respondents who are library directors (71.3%) is smaller than the percentage for NPs. Contacts at POs passed questionnaires to library staff (13.8%), CIOs (6.2%), and archivists (3.7%). Of the four people classed in "Other," two hold combined positions (i.e., library director-CIO and library director-archivist), one is assistant director of administrative services, and one is a faculty member affiliated with the library.

At PPT institutions, the percentage of respondents who are library directors (48.3%) is smaller than the percentages for POs and NPs. Contacts at PPTs mostly passed questionnaires to associate or assistant library directors (26.7%), library staff (13.3%), and archivists (3.3%). They rarely passed questionnaires to CIOs (1.7%). The four people classed in "Other" hold different positions: two are digital library directors, one is the associate dean for research, and one is the assistant director of library campus support systems.

At IMP institutions, the percentage of respondents who are library directors drops to 21.6%. Larger percentages of library staff generally (43.3%) and associate or assistant library directors (27.0%) are completing questionnaires. The two people classed in “Other” are a director of special collections and archives and a librarian working as a temporary consultant on the IR.

As shown in Table 2.1, library-director percentages decrease as IR activity increases. Most likely, directors passed on our request to complete questionnaires to staff who would be more knowledgeable about IR activity at their institution. Such activity increasingly involves library staff generally at PO institutions, and both library staff generally and assistant and associate directors at PPT and IMP institutions.

2.2 Using CCHE to Characterize Participating Institutions

Scrutinizing the types of institutions that participated in the MIRACLE census respondents made project investigators wonder whether certain types of institutions are more or less likely to be involved with IRs. To characterize the institutions that participated in the MIRACLE Project, investigators turned to the Carnegie Classification of Institutions of Higher Education (CCHE). CCHE was derived from empirical data on colleges and universities and has been updated five times since it was originally published for use by researchers in 1973. CCHE is “the leading framework for describing institutional diversity in U.S. higher education [and] ... has been widely used in the study of higher education, both as a way to represent and control for institutional differences, and also in the design of research studies to ensure adequate representation of sampled institutions, students, or faculty” (Carnegie Foundation 2006a).

Table 2.2 lists classes of CCHE institutions that MIRACLE Project investigators invited to participate in the nationwide census. Because we limited the census to institutions awarding four-year baccalaureate degrees or higher, missing from Table 2.2 is the CCHE “Associate” class for institutions awarding two-year associate degrees.

To emphasize the *research*-intensive nature of census institutions, MIRACLE Project staff broke up the “Doctorate-granting Universities” CCHE class into two classes: (1) “Research universities,” bearing institutions from the two Research Universities (RU/VH and RU/H) CCHE subclasses; and (2) “Doctoral universities,” bearing institutions from the “Doctoral” (DRU) CCHE subclass. Figure 2.3 shows percentages of institutions participating in the MIRACLE Project census by CCHE classes. It also distributes the population of U.S. academic institutions into these same CCHE classes.

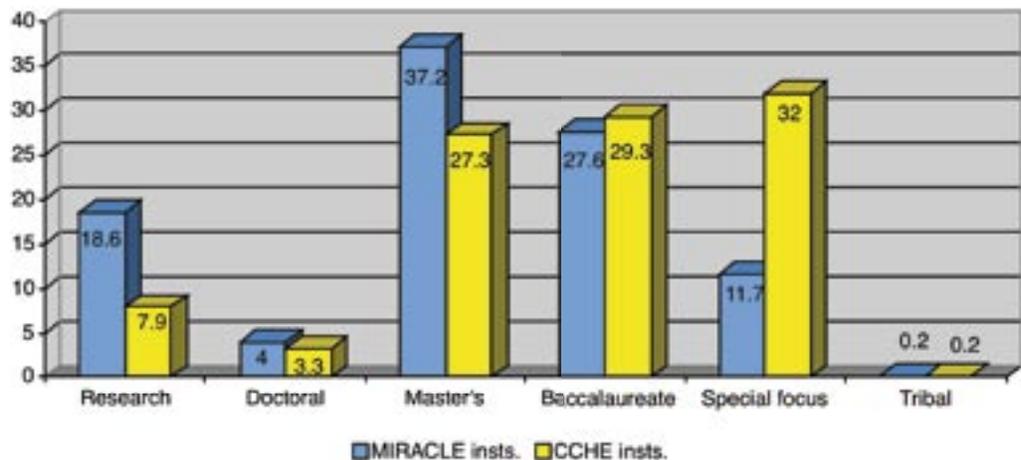
Table 2.2. Classes of CCHE institutions invited to participate in the MIRACLE census

Class	Definition	Subclasses
Doctorate-granting universities*	Institutions that award at least 20 doctoral degrees per year (excluding doctoral-level degrees such as the JD, MD, PharmD, DPT, which qualify recipients for entry into professional practice).	RU/VH: Research Univs. (very high research activity) RU/H: Research Univs. (high research activity) DRU: Doctoral/Research Univs.
Master's colleges and universities*	Institutions that award at least 50 master's degrees per year.	Master's/L: Master's Colleges and Univs. (larger programs) Master's/M: Master's Colleges and Univs. (medium programs) Master's/S: Master's Colleges and Univs. (smaller programs)
Baccalaureate colleges*	Institutions where baccalaureate degrees represent at least 10% of all undergraduate degrees and that award fewer than 50 master's degrees or fewer than 20 doctoral degrees per year.	Bac/A&S: Baccalaureate Colleges—Arts and Sciences Bac/Diverse: Baccalaureate Colleges—Diverse Fields Bac/Assoc: Baccalaureate/Associate's College
Special-focus institutions	Institutions awarding baccalaureate or higher-level degrees where a high concentration of degrees is in a single field or set of related fields.	Examples are theological seminaries, Bible colleges, medical schools, engineering schools, business schools, and law schools.
Tribal schools	Colleges and universities that are members of the American Indian Higher Education Consortium.	The majority are associate's colleges but there are a few baccalaureate colleges.

* Excludes special-focus institutions and tribal colleges.

High percentages of CCHE institutions come from the Special Focus (32.0%), Baccalaureate (29.3%), and Master's (27.3%) CCHE classes. Except for special-focus (11.7%) institutions, percentages of MIRACLE census respondents are comparable for master's (37.2%) and baccalaureate (27.6%) institutions. A large percentage (18.6%) of MIRACLE census respondents are research universities, but only 7.9% of CCHE institutions come from this class.

Figure 2.3. MIRACLE institutions versus CCHE institutions



MIRACLE Project staff tallied the four types of MIRACLE census respondents according to their CCHE class. The results (see Table 2.3) reveal what types of CCHE institutions are more and less likely to implement IRs.

Research universities vastly outnumber all other CCHE classes involved with IMP and PPT. A few institutions in the other CCHE classes are implementing IRs, but most are in the PO stage or are not involved with IRs at all. Most NP and PO respondents come from master's and baccalaureate institutions.

Table 2.3. CCHE classes reveal the extent of IR involvement by census respondents

CCHE classes	NP		PO		PPT		IMP		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
Research univs.	13	5.5	14	15.2	26	37.1	30	62.5	83	18.6
Doctoral univs.	7	3.0	7	7.6	3	4.3	1	2.1	18	4.0
Master's	103	43.6	32	34.8	22	31.5	9	18.8	166	37.2
Baccalaureate	79	33.5	29	31.5	10	14.3	5	10.4	123	27.6
Special focus	32	13.6	9	9.8	8	11.4	3	6.2	52	11.7
Tribal	1	0.4	0	0.0	0	0.0	0	0.0	1	0.2
Unclassified*	1	0.4	1	1.1	1	1.4	0	0.0	3	0.7
Total	236	100.0	92	100.0	70	100.0	48	100.0	446	100.0

* Three institutions are unclassified because they responded with two or more institutions in partnership-like arrangements.

2.3 Positions of the People Involved with IRs

Questionnaires asked census respondents about the people involved in their institutions' IR efforts. Specifically, respondents from PO, PPT, and IMP institutions were asked, "How active were people in the following positions in terms of leading the charge to get involved with IRs at your institution?" and respondents from NP institutions were asked, "How active do you think that the people in these positions would have to be to light the spark for IR planning at your institution?" Respondents chose from a list of 13 positions or could write in a response for "Other."

To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very active; (+1) somewhat active; (0) no opinion, don't know, or not applicable; (-1) somewhat inactive; and (-2) very inactive. They totaled the weights. The results were then compiled to rank order all the positions. Table 2.4 uses IMP ranks to order top- (1 to 4), middle- (5 to 8), and bottom-ranked (9 to 13) positions.

Generally, PO, PPT, and IMP respondents agree about top-, middle-, and bottom-ranked positions. For PO, PPT, and IMP respondents, the three top-ranked active positions are (1) library director, (2) assistant or associate library director(s), and (3) library staff member(s). Students at all levels, faculty governance, presidents, and chancellors are not necessarily active in IR efforts.

Table 2.4. Positions of people involved in the IR effort

Top-ranked positions (1 to 4)	NP	PO	PPT	IMP
Library director	1	1	1	1
Assistant library director(s)	(11)†	2T*	2	2
Library staff member(s)	(5)	2T	3	3
A particular faculty member	(10)	(6)	4	4
Middle-ranked positions (5 to 8)	NP	PO	PPT	IMP
Institution's archivist	7	(4)	5	5
Institution's vice president or provost	(2)	8T	(9)	6
Staff at a library network, consortium, or other affiliated group	6	5	6	7
Faculty members generally	(3)	8T	8	8
Bottom-ranked positions (9 to 13)	NP	PO	PPT	IMP
Institution's chief information officer	(4)	(7)	(7)	9
Faculty governance	(8)	12	12	10
Graduate students	12	10	10	11
Institution's president or chancellor	9	13	13	12
Undergraduate students	13	11	11	13

† Parentheses indicate NP, PO, and PPT positions that deviated from IMP top, middle, or bottom ranks.

* T indicates a ranked position that tied another position's weight.

NP respondents agree with those in the other three groups that the highest level of activity comes from the library director. To light the spark for the IR effort at NP institutions, support must come not only from top positions in the library but also from other leaders at the institution (e.g., the vice president or provost, or CIO) and from faculty members generally.

The questionnaires allowed respondents to give open-ended responses to this question. Most responses are unique but a few overlap. Respondents from PO, PPT, and IMP institutions say the following people are very active:

- director, staff, and/or advisory committee from the institution's instructional technology unit (three responses)
- faculty at the institution's Office of Undergraduate Research Initiatives
- global subject discipline research committees
- internal and external volunteers such as library school students and visiting librarians
- public relations staff
- strategist from the institution's Academic Computing
- curriculum technology staff
- technology-assisted learning staff

At NP institutions, people in positions that would have to be very active in an IR effort are again external to the library:

- director of the institution's instructional technology unit
- academic deans
- museum director
- Web developers
- systems staff

An examination of open-ended responses reveals in retrospect that we should have included a response category for “instructional technology staff” because it might have figured among the top-ranked response categories at PO, PPT, and IMP institutions. Other response categories on our list should have been “academic computing” and “academic deans.”

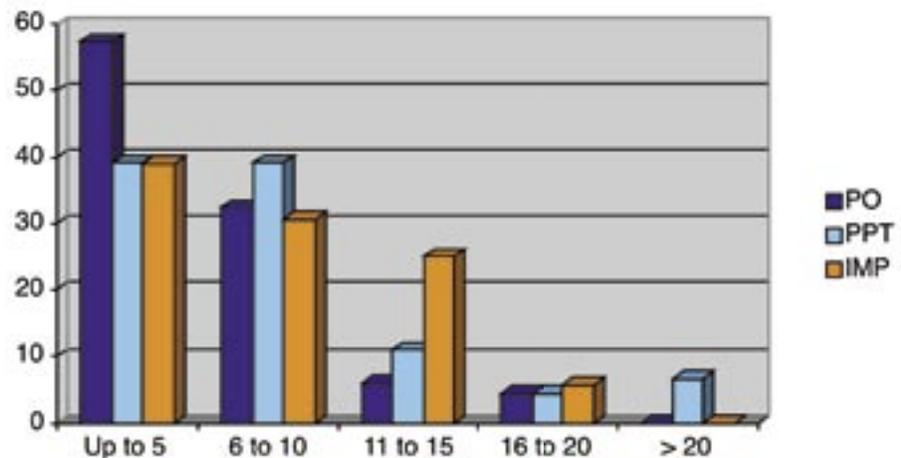
2.4 Number of People Involved with IRs

Questionnaires asked PO, PPT, and IMP respondents how many people were involved in their institutions’ IR efforts. The overall average is 7.2 people. At PO, PPT, and IMP institutions, averages are 6.0, 8.4, and 7.8 people, respectively.

These are merely counts of the number of people involved in the IR effort. MIRACLE Project investigators also wanted to ask census respondents about full-time equivalents (FTEs). However, respondents who pretested MIRACLE questionnaires expressed difficulty generating exact FTE numbers so we deleted questions about FTEs.

Figure 2.4 presents results in five-person ranges. Up to 10 people is typical for 89.8% institutions in the PO stage. On average, when institutions move to the PPT stage, the number of people involved increases. It then decreases in the IMP stage. At some institutions, more than 20 people are involved in the IR effort at the PPT stage. Although numbers from the ARL SPEC Kit are substantially higher (Bailey et al. 2006, 15), both SPEC Kit numbers and our numbers show a downward trend between the PPT and the IMP stages (see Appendix F4).

Figure 2.4. Number of people involved in the IR effort



2.5 Positions of People Involved with IRs

The questionnaires asked who is leading IR planning, planning and pilot testing, and implementation at PO, PPT, and IMP institutions. Table 2.5 gives the results.

Table 2.5. Positions of people leading the IR effort

Position leading the IR effort	PO		PPT		IMP	
	No.	%	No.	%	No.	%
Library director	46	54.7	18	28.6	13	31.7
Library staff member	12	14.3	15	23.8	14	34.2
Assistant or associate library director	3	3.6	13	20.6	11	26.8
CIO	4	4.8	1	1.6	1	2.4
Archivist	5	5.9	2	3.2	0	0.0
Faculty member in an academic unit	4	4.8	1	1.6	0	0.0
No committee or committee chair has been charged	6	7.1	1	1.6	0	0.0
Other	4*	4.8	12†	19.0	2‡	4.9
Total	84	100.0	63	100.0	41	100.0

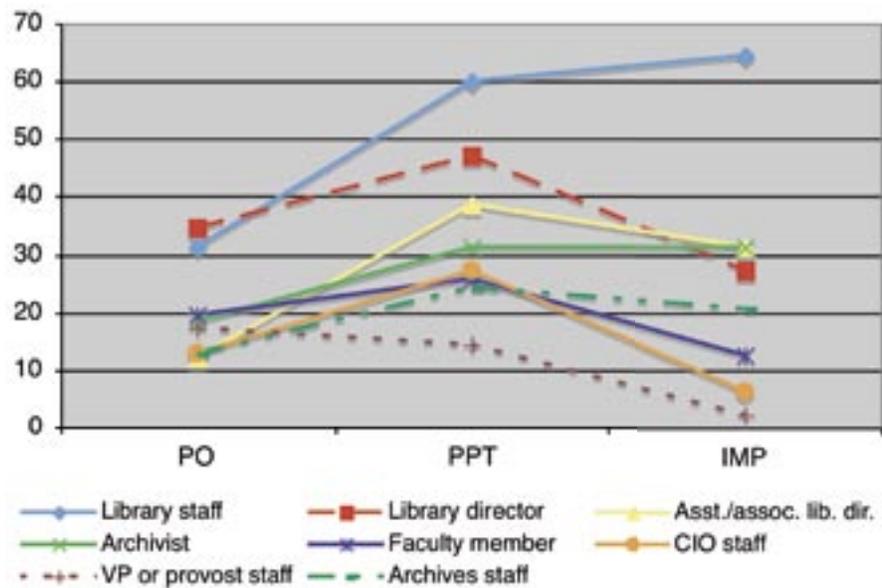
* Team effort (2); consortium (1); duo effort: archivist and library director (1).

† Director of instructional technology (3); duo effort (3), for example, library director and CIO; vice president or associate dean for research (2); team effort (2); consortium (1); digital asset management committee (1).

‡ Consortium (1); director of special collection and archives (1).

Generally, people in *library* positions lead in all stages of the IR effort. In the PO stage, the library director is in the lead. The library director relinquishes that role when the IR effort reaches PPT and IMP stages, in most cases, to one particular staff member or an assistant or associate library director. If archivists, CIOs, and faculty members from academic units are in the lead, the IR effort is in the planning stage. Write-in responses reveal that staff from two or more units may share the lead, especially during PPT stage. For example, the associate librarian leads planning and the CIO leads pilot testing.

Figure 2.5. IR committee membership



Questionnaires for PO, PPT, and IMP institutions asked respondents what positions IR committee members held. Figure 2.5 depicts the percentages of respondents choosing from 13 positions listed on their questionnaires. The figure uses lines to connect the three percentages, beginning with PO and ending IMP. It is not a timeline because different people completed PO, PPT, and IMP questionnaires. Presenting results in this way is helpful because it reveals the following dynamics of committee membership:

- Generally, IR committees are more inclusive during the PPT stage and less inclusive during the PO and IMP stages.
- The likelihood that staff from the vice president's or provost's office are on IR committees *decreases* from the PO stage to the PPT stage, while people in all other positions are *more* likely to be members of IR committees.
- The likelihood that library staff and assistant or associate library directors are on IR committees *increases* from stage to stage, while people in all other positions are *less* likely to be members of IR committees as IR work continues.
- Faculty members are *more* likely to be involved in the conceptual stages of planning the IR; their involvement *decreases* as the IR becomes operational.

Especially nonrepresentative at the IMP stage are CIO staff, faculty members, and staff from the office of the vice president or provost. Not included in Figure 2.5 are percentages for five positions—graduate students, undergraduate students, and staff from the office of the president or chancellor, from the CIO's office, and from the institution's legal counsel—because less than 10% of respondents observed their participation on IR committees.

Many respondents wrote in unique staff or management positions not included in the list.

At PO institutions the write-in positions are

- three for academic computing and three for instructional technology
- two for consortium
- one each for alumni relations, art-slide curator, center for teaching, communications, development, enrollment services, external affairs, dean of graduate studies, staff photographer, student affairs, and university press

At PPT institutions the write-in positions are

- four for instructional technology
- two for consortium
- one each for academic computing, art-collections curator, art-slide curator, and dean of graduate studies

At IMP institutions the write-in positions are

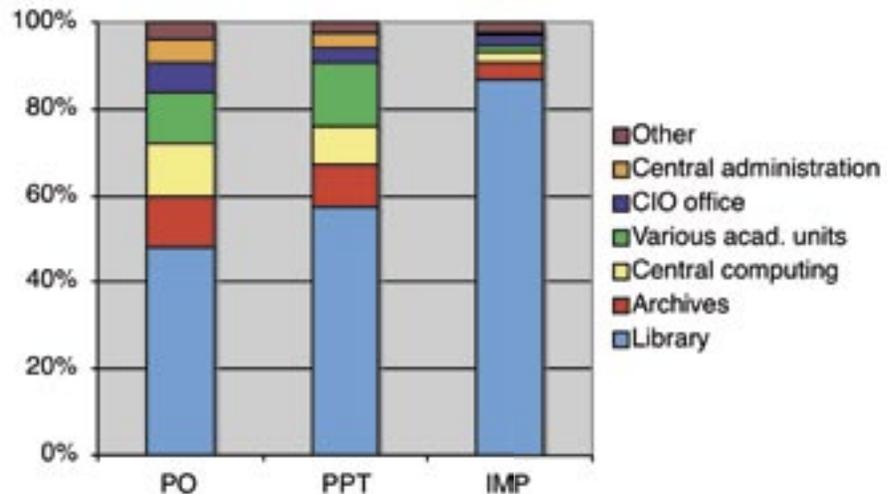
- three for instructional technology
- one each for academic computing, digital library program, health sciences center, media services, university press, and the college-level Web content editor

In retrospect, we should have included response categories for “instructional technology” and “academic computing” because several respondents volunteered them.

2.6 The Responsibility for the IR

Questionnaires for IMP and PPT respondents asked what percentage of the responsibility for an operational IR is given to various campus units, and the questionnaire for PO respondents asked what percentage should be given to various campus units. All three questionnaire versions listed the same units and respondents could write in units that were not listed. MIRACLE Project staff programmed Survey-Monkey to force respondents to enter percentages that added to 100%. Figure 2.6 gives the results.

Figure 2.6. Responsibility for the IR



During planning, respondents share responsibility for the IR, with the library taking about 40% of the responsibility; archives, central computing, and various academic units, sharing about 12% of the responsibility; and the CIO's office sharing 6% of the responsibility. During PPT, the responsibilities of the library and various academic units increase while others' responsibilities decrease. The increase for academic units during the PPT phase probably entails early adopters who are contributing to the IR in a pilot test. During IMP, the library shoulders almost all of the responsibility for the IR. Questionnaires should have included a response category for “consortium” because several write-ins named their consortium as the unit taking most of the responsibility for the IR.

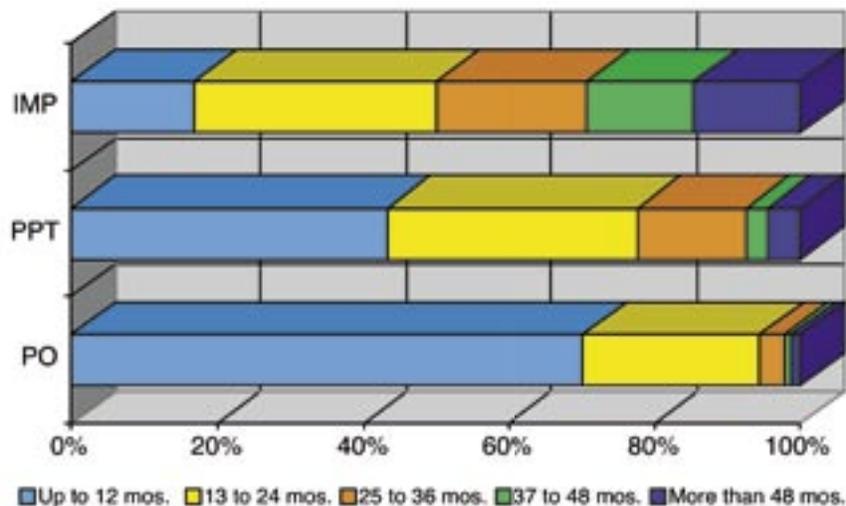
2.7 Involvement with IRs

Questionnaires asked respondents at PO, PPT, and IMP institutions how long they have been involved with IRs. PO institutions average 12 months, PPT institutions average 21.3 months, and IMP institu-

tions average 31.5 months.

Figure 2.7 shows responses in 12-month ranges. About 70% of PO institutions have been involved with IRs for 12 months or fewer. Comparable percentages of PPT (77.6%) and IMP (70.8%) institutions have been involved with IRs for 24 or fewer months and 36 or fewer months, respectively. About 15% of IMP institutions have been involved with IRs for more than four years.

Figure 2.7. Involvement with IRs



2.8 Chapter 2 Summary

Of the 2,147 academic library directors or senior library administrators MIRACLE Project staff contacted, 446 participated in the census—a response rate of 20.8%. A little more than half of respondents have done no IR planning to date, about 20% are planning for IRs, about 15% are actively planning and pilot testing IRs, and a little more than 10% have implemented an operational IR (see Figure 2.1).

MIRACLE Project investigators used the CCHE to determine the types of institutions that are more or less likely to be involved with IRs. “Research universities” vastly outnumber all other CCHE classes involved with IMP and PPT (see Table 2.3). Most NP and PO respondents come from master’s and baccalaureate institutions.

Census respondents in the PO, PPT, and IMP stages agree on the positions of people most involved with IRs at their institution. They are the library director, assistant or associate library director(s), and library staff member(s) (see Table 2.4). To light the spark for the IR effort at NP institutions, support must come not only from the library director but also from other leaders at the institution, including the vice president or provost and CIO. Faculty members generally should also be active.

The number of people involved in the IR effort averages 7.2 overall but varies a little during the IR implementation process. PO, PPT, and IMP institutions average 6.0, 8.4, and 7.8 people, respectively (see Subchapter 2.4). The PPT stage is most inclusive, involving 20

or more people at times (see Figure 2.4).

In terms of the person leading the IR effort, the library director takes the lead in the planning stage but relinquishes it, in most cases, to one particular staff member or an assistant-associate library director in the PPT and IMP stages (see Table 2.5). If archivists, CIOs, and faculty members from academic units are in the lead, the IR effort is probably in the planning stage.

IR committee membership waxes and wanes depending on the particular phase of the IR project (see Figure 2.5). IR committees are most inclusive during the PPT stage and less inclusive during the PO and IMP stages. The likelihood that library staff and assistant or associate library directors are on IR committees *increases* from stage to stage while people in all other positions are *less* likely to be members of IR committees as work proceeds.

During planning, respondents share responsibility for the IR with the library taking about 40% of the responsibility; archives, central computing, and various academic units, sharing about 12% of the responsibility; and the CIO's office sharing 6% of the responsibility (see Figure 2.6). When planning and pilot testing, the responsibilities of the library and various academic units increase while others' responsibilities decrease. The increase for academic units during the PPT phase probably entails early adopters who are contributing to the IR in a pilot test. At implementation, the library shoulders almost all of the responsibility for the IR.

Asked how long their institutions have been involved with IRs, PO institutions average 12 months, PPT institutions average 21.3 months, and IMP institutions average 31.5 months (see Figure 2.7). Of the total 48 IMP institutions in the MIRACLE census, seven (14.6%) have been involved with IRs for more than four years.

3 THE BUDGET FOR AN IR

Chapter 3 focuses on the budget for an institutional repository (IR), specifically on sources of funding and on line items in the IR budget.

3.1 Sources of Funding

Respondents from institutions planning (PO), planning and pilot testing (PPT), and implementing (IMP) IRs were asked how likely the funding for an IR was to come from a list of 17 sources.

To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very likely; (+1) somewhat likely; (0) no opinion, don't know, or not applicable; (-1) somewhat unlikely; and (-2) very unlikely. Staff totaled the weights. These results were compiled to rank order all the funding sources. Table 3.1 uses IMP ranks to order the top- (1 to 6), middle- (7 to 12), and bottom-ranked (13 to 17) funding sources.

Table 3.1. Top- and bottom-ranked funding sources

Top-ranked funding sources (1 to 6)	PO	PPT	IMP
Special initiative supported by the library	1	2	1
Costs absorbed in routine library operating costs	2	1	2
Regular budget line item for your institution's library	4	3	3
Grant awarded by an external source	3	4	4
Special initiative supported by your institution's central administration	5	6	5
Special initiative supported by your institution's archives	(8)†	(9)	6
Middle-ranked funding sources (7 to 12)	PO	PPT	IMP
Grant awarded by an internal source	11	11	7
Special initiative supported by your institution's central computer services	(6)	(5)	8
Regular budget line item for your institution's archives	(15)	10	9
Costs absorbed in routine operating costs of your institution's archives	13	8	10
Regular budget line item for your institution's central computer services	9	11	11
Regular budget line item for your institution's central administration	(7)	(15)	12
Bottom-ranked funding sources (13 to 17)	PO	PPT	IMP
Costs absorbed in routine operating costs of your institution's central computer services	(10)	(7)	13
Costs absorbed in routine operation costs of central administration	(7)	15	14T*
Special initiative supported by academic colleges, departments, and schools	(11)	(13)	14T
Costs absorbed in routine operating costs of academic colleges, departments, and schools	16	16	16
Regular budget line item for academic colleges, departments, and schools	17	17	17

† Parentheses indicate PO and PPT funding sources that deviated from IMP top, middle, or bottom ranks.

* T indicates a ranked funding source that tied another source's weight.

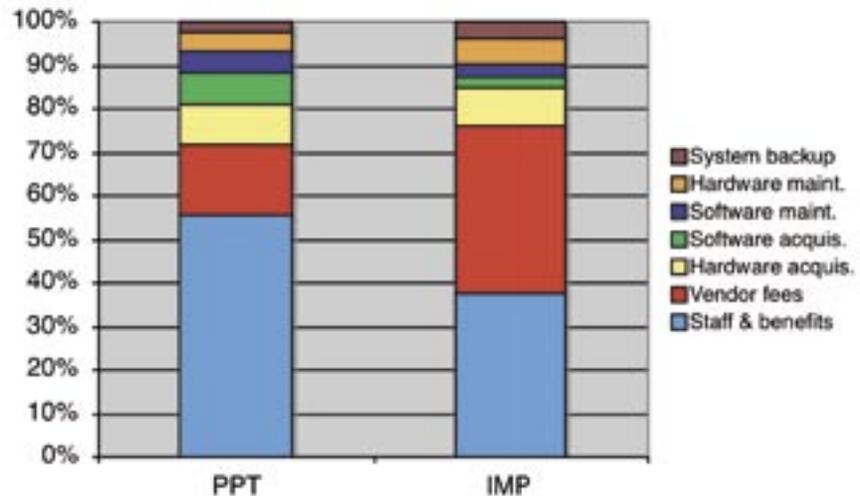
PO, PPT, and IMP respondents agree about the top-ranked funding source for IRs—funding comes or will come from the *library*. For example, a typical strategy is to absorb the costs into routine library operating expenses. Libraries at large research universities may find it easier to enlist such a strategy than libraries at small institutions because the IR effort may be hard to pinpoint in the former's multimillion-dollar budgets. Other strategies, such as a special library initiative or adding a regular budget line item, may require the library to obtain support from the central administration or to divert resources from ongoing activity to the IR.

Respondents agree that IR funding does not or will not come from academic units. Respondents from PO institutions especially do not envision funding coming from the archives. In write-in responses, two institutions indicate that they have received, or expect to receive, funds from the U.S. Department of Education's Title III grants, which aim to "assist eligible IHEs [institutions of higher education] to become self-sufficient and expand their capacity to serve low-income students by providing funds to improve and strengthen the academic quality, institutional management, and fiscal stability of eligible institutions" (U.S. Department of Education 2006). Two other write-in responses say their funding comes from a consortium.

3.2 IR Budget Line Items

Questionnaires for IMP and PPT institutions asked respondents what percentage of their IRs' annual budget is allocated to various line items. The two questionnaire versions listed the same line items, and respondents could write in items that were not listed. SurveyMonkey was programmed so that it required respondents to enter percentages that added to 100. Figure 3.1 gives the results.

Figure 3.1. IR budget line items



Costs for staff and vendor fees represent about 75% of the IR budget, with staff costs exceeding vendor fees during PPT and vice versa during implementation. Hardware acquisition represents about 10% of the IR budget. Software costs represent 7% and 2.5% of PPT and IMP budgets, respectively. Costs for software maintenance, hardware maintenance, and system backup account for 12.5% of the IR budget.

Although many respondents volunteered open-ended comments pertaining to costs, only two comments cite line items that MIRACLE Project investigators failed to include in their original list:

- "Marketing and PR [public relations] activities."
- "Server farm charges, 1%, hosted by Central Computing; storage farm charges, 11%; indirect costs (15%)."

Despite having a fully functional and operational IR, several IMP respondents write about the informality of their IR's budget:

- "Our IR does not have a budget."
- "No specific budget for IR. It is absorbed in the library budget."
- "We do not really have a budget for this. The fee to the vendor is paid out of our library's operating costs. Three staff members each spend a few hours a week working on this project. It is impossible to estimate the staff cost."
- "We only budget for the subscription to our hosted product. We don't budget the staff time."
- "IR isn't budgeted separately anymore and was only partially budgeted separately from the library in year 1 and year 2."

- “[T]his question is difficult to answer. Staff responsible for the repository [are] doing repository work [and other unrelated tasks]. Vendor fees are shared between the library and central administration . . . I don’t know the full operating budget of [either the library or central administration] nor am I interested to know.”

Here is a comment from an IMP respondent who is exceptionally precise about her institution’s IR budget:

- “This coming year will be an exception: \$100,000 has been allocated for initial purchase and migration of commercial service provider. Our operating budget alone without the above would be staff, 49%; hardware acquisition, 27%; hardware maintenance, 3%; software acquisition, 16%; software maintenance and updates, 2%; vendor fees, 3%.”

Several PPT respondents comment on the shared nature of the IR initiative:

- “We are not funding this project with dollars from our [library] budget; system administration is picking up all hardware and software costs. We [the library] are providing only human resources.”
- “Our IR software and hardware were a special allocation from Instructional Technology Services (ITS) and the central administration. Maintenance and upgrade of server and software will be absorbed by ITS regular budget process. Implementation of the IR will be absorbed into regular library workflow.”
- “The software license of ContentDM was purchased by central computing. The annual maintenance license agreement is paid by the library. All labor is carved from staff time in the Library and Institutional Technology Departments, with faculty involvement supervising work-study students. We are small scale, concentrating on unique content when faculty want something digitized.”
- “Our IR is distributed among departments on campus—it has no separate budget.”

Other PPT respondents could not break down IR costs into listed line items because they did not know or were unsure about the breakdown, or had not yet budgeted for the IR.

3.3 Chapter 3 Summary

PO, PPT, and IMP respondents agree about the top-ranked funding sources for IRs—funding comes or will come from the *library* (see Table 3.1). They also agree that funding is *not* coming from academic units.

Costs for staff and vendor fees represent about 75% of the IR budget, with staff costs exceeding vendor fees during PPT and vice versa during implementation (see Figure 3.1). Hardware acquisition represents about 10% of the IR budget. Software costs represent 7% of PPT and 2.5% of IMP budgets. Costs for software maintenance, hardware maintenance, and system budget account for 12.5% of the

IR budget. Underlying the write-in responses of several IMP respondents is a certain informality about the IR budget. We did not speculate on reasons for this informality.

4 IMPORTANT INVESTIGATIVE ACTIVITIES

Chapter 4 explores important investigative activities that institutions conduct to determine whether to implement an institutional repository (IR).

4.1 Important Investigative Activities

Planning only (PO), planning and pilot testing (PPT), and implementation (IMP) questionnaires asked respondents to rate the importance of various investigative activities in terms of influencing their decision to initiate planning, pilot testing, and implementation. To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very important; (+1) somewhat important; (0) no opinion, don't know, or not applicable; (-1) somewhat unimportant; and (-2) very unimportant. They totaled the weights. These results were then compiled to rank order all the activities. Table 4.1 uses ranked activities in the "Total" column to order top-, middle-, and bottom-ranked activities.

Table 4.1. Important investigative activities

Top-ranked investigative activities (1 to 4)	PO	PPT	IMP
Learning about successful implementations at comparable institutions	1	2	1
Learning from reports of other institutions' PO, PPT, and IMP activities	2	1	2
Learning about successful implementations at a wide range of academic institutions	(5)*	3	3
An analysis of a thorough literature review of IRs	(9)	(5)	4
Middle-ranked investigative activities (5 to 8)	PO	PPT	IMP
Using other institutions' operational IRs	6	8	5
Results of your institution's needs assessment	7	7	6
Demonstrating operational IRs to my institution's decision makers	(3)	6	7
Learning about available expertise and assistance from a library consortium, network, group of libraries, etc.	(4)	(4)	8
Bottom-ranked investigative activities (9 to 12)	PO	PPT	IMP
Demonstrating IR metadata harvesters such as OAIster and Google Scholar to my institution's decision makers	10	10	9
Identifying better digital preservation techniques	(8)	9	10
Waiting for a critical mass of IR implementation at comparable institutions to happen	11	12	11
Waiting for a critical mass of IR implementation generally to happen	12	11	12

* Parentheses indicate PO and PPT investigative activities that deviated from IMP top, middle, or bottom ranks.

At the top of the ranked list are investigative activities concerning learning about IRs from the experiences of others. For PPT and IMP respondents, this includes analyzing literature reviews. PO and PPT respondents rank “Learning from a library consortium ...” higher than IMP respondents do, most likely because the latter charged ahead with IR implementation before consortia, networks, and comparable groups had begun their involvement with IRs. PO respondents rank “Demonstrating operational IRs to my institution’s decision makers” much higher than PPTs and IMPs do. Such demonstrations probably make IRs more tangible to decision makers. They increase decision makers’ understanding of system functionality, IR contributors, contents, users, and uses. They help decision makers understand how IRs are in keeping with the institution’s mission and thereby make them more favorably inclined to the IR initiative in terms of both funding and rhetoric.

At the bottom of the list are two activities about “Waiting for a critical mass of IR implementation to happen.” Because only 28% of both PO and PPT respondents rate it “very” or “somewhat” important, it is clear that these respondents want to get involved with IRs now rather than to follow the crowd.

A wait-and-see attitude is evident in this write-in comment:

- “Waiting for clear leaders to emerge in the vendor or open-source IR options. Waiting for options that better meet our needs. Many products have potential but [are] not ready for prime time.”

Write-ins by several PO respondents reveal three investigative activities that, had they been listed on the questionnaire, would have received high ratings: (1) finding funding for IR hardware and software; (2) finding funding for IR staffing; and (3) finding expertise for IR staffing.

Two write-ins by PPT respondents describe how they are taking the initiative to study their institutions’ digital output:

- “IRs were starting to be formed on an ad hoc basis across the campus; we wanted to provide a single gathering space and search engine for these documents.”
- “Conducted study of [our] institution’s Web presence, which demonstrated a stewardship need and identified an extensive amount of potential IR content. Worked with pilot departments to add content and gauge interest.”

Two write-ins by IMP respondents capture of experience of early adopters of IR technology:

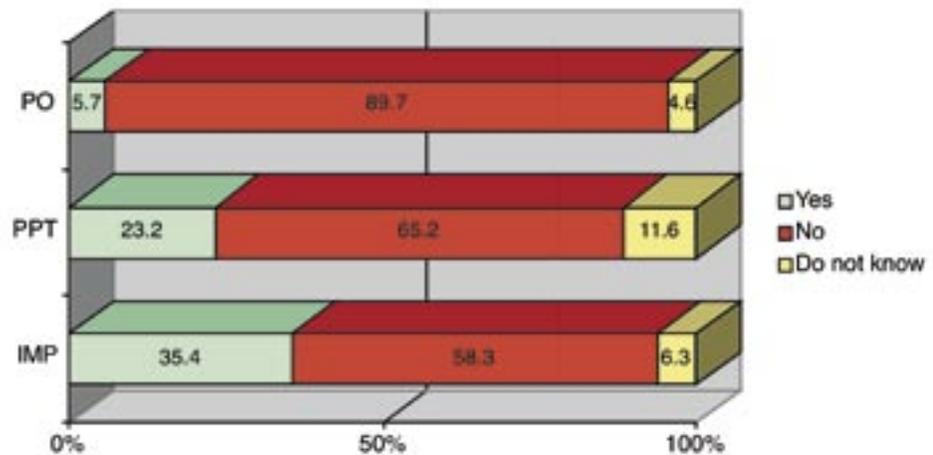
- “We agreed to become a member of the original DSpace Federation in order to test a repository system and position ourselves to engage in e-publishing activities.”
- “We were very early in our implementation, so there were few fully implemented repositories to examine. It was our provost’s desire to start a ‘faculty e-archive’ that was the primary deciding factor.”

4.2 The Needs Assessment

MIRACLE Project investigators expected that a needs assessment would be an important investigative activity that institutions would undertake before deciding to get involved with IRs. For that reason, the questionnaires featured as many as three additional questions about the needs assessment.

To our surprise, respondents ranked the needs assessment in the middle of the pack (Table 4.1). Although most evidently felt that the needs assessment was relatively unimportant compared with other activities, their answers were revealing.

Figure 4.1. Did your institution conduct a needs assessment?



About one in sixteen PO, one in four PPT, and one in three IMP institutions, respectively, have conducted a needs assessment (Figure 4.1). The percentage of respondents who do not know whether their institution conducted a needs assessment ranges from 5% to 12%. Asked whether they would be likely to conduct a needs assessment prior to making a decision about implementing an IR, about 70% of PO respondents and 44% of PPT respondents say they are “very” or “somewhat” likely to do so.

Questionnaires asked IMP respondents how important the needs assessment was for accomplishing 11 IR-related tasks. Table 4.2 lists these tasks and the percentages of respondents who told us that the needs assessment is “very” or “somewhat” important for accomplishing them. More than 75% of respondents rate all but four tasks very high in importance. At the top of that list is “Formulating IR policies.” Because “Making the decision to implement an IR” is close to the bottom of the list, we can presume that census institutions were not conducting the needs assessment to help them decide whether to implement an IR. Instead, they were conducting the assessment to discover the reception their IR would get from their institution’s learning community. One write-in comment says as much:

- “This was not a traditional needs assessment. We knew we were going to implement an IR and some of the needs assessment was carried out while planning the IR.”

Table 4.2. Importance of the needs assessment

Rank	IR-related tasks	% Important
1	Formulating IR policies	90.0
2	Identifying first adopters of an IR	84.2
3	Recruiting digital content for the IR	83.3
4	Choosing an IR software package	82.4
5	Streamlining IR planning and implementation	82.4
6	Increasing faculty awareness of the IR	79.0
7	Identifying especially active contributors to the IR	77.8
8	Identifying new services to build onto the IR	72.2
9	Scheduling the rollout of various IR services	68.8
10	Making the decision to implement an IR	68.4
11	Identifying preservation techniques	62.5

A handful of respondents told us that faculty interest was key to proceeding with an IR and that they did not necessarily have to conduct a needs assessment to find signs of such interest. Here are their comments in this regard:

- “There was no needs assessment but the IR was very much faculty driven. Leadership was taken by the University Library Council (a senate-provostial advisory group) that pushed the agenda and prepared the report that led to provost funding and support.”
- “Our assessment was more dynamic and ongoing ... it involved response to innovative faculty requests and ongoing outreach from librarians regarding changes in scholarly communication practices, e.g., an e-publishing symposium hosted by the library author’s rights issues.”
- “Our former dean of faculty was particularly interested in DSpace and secured funding for the university libraries to implement and support its use here.”

4.3 Pilot Testing IR Software Packages

The PPT and IMP questionnaires asked respondents to rate the importance of various benefits of pilot testing one or more IR-system software packages. To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very important; (+1) somewhat important; (0) no opinion, don’t know, or not applicable; (-1) somewhat unimportant; and (-2) very unimportant. They totaled the weights. These results were then compiled to rank order all the positions. Table 4.3 lists all 10 benefits in rank order. Except for the bottom-ranked benefit, the percentages of respondents rating benefits “very” or “somewhat” important are very high, ranging from 67% to 93%. Respondents are positive even about the bottom-ranked benefit, giving demonstrations to prospective partners, because almost 50% of them rate it “very” or “somewhat” important.

Table 4.3. Important benefits of pilot testing

Important benefits (1 to 5)	PPT	IMP
Identifying the strengths and shortcomings of available IR software	2	1
Developing the requisite technical expertise for IR implementation	1	2
Estimating costs for the technical implementation of an IR	3T*	3
Giving demonstrations to people involved in the IR implementation decision	5	4
Identifying first adopters of an IR at your institution	(6)	5
Less important benefits (6 to 10)	PPT	IMP
Preservation of your institution's intellectual output	(3T)	6
Gauging the interest of potential contributors to the IR	7	7
Control over your institution's intellectual output	9	8
Gauging the interest of potential IR-system users	8	9
Giving demonstrations to an institution(s) interested in partnering with us to encourage them in IR implementation	10	10

† Parentheses indicate PPT benefits that deviated from IMP top- and bottom-ranked benefits.

* T indicates a ranked benefit that tied another benefit's weight.

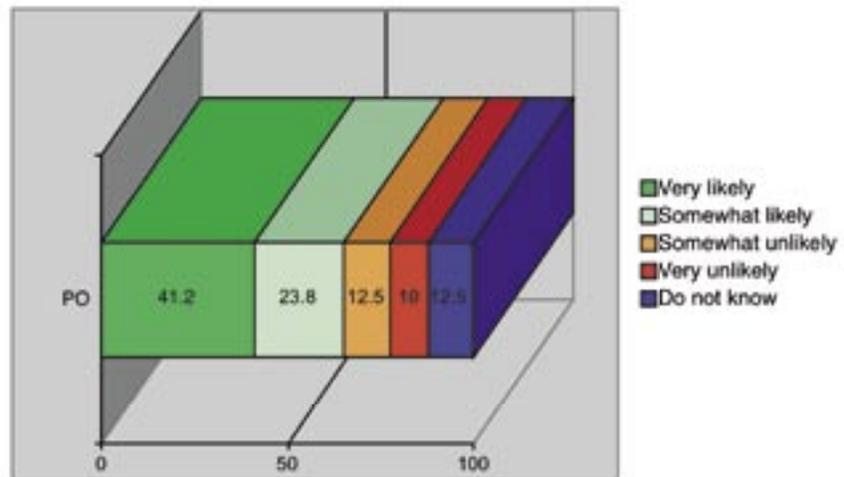
The three top-ranked benefits—developing the requisite technical expertise, learning about IR software, and estimating costs—are very practical in terms of implementing an IR. Middle- to low-middle ranked benefits pertain to potential contributors and users of the IR. MIRACLE Project investigators thought benefits pertaining to contributors especially would be ranked higher in view of the difficulty in recruiting each (see Subchapter 6.5 and Appendix F8.4), but they were not. PPT and IMP respondents' lists of ranked benefits are almost the same. The only difference is that PPT respondents give greater importance to preserving their institution's intellectual output. Two write-ins comment on the importance of pilot testing for collection building:

- "Building an IR collection prior to production so that on [the] release [of our operational IR] it has apparent value."
- "Expanding student access to teaching materials in particular courses such as archaeology and botany."

Questionnaires asked PO respondents whether they were likely to pilot test one or more IR software packages prior to implementing an IR. Figure 4.2 graphs the results.

Almost two-thirds of PO respondents are likely to pilot test. Whether the one-quarter of PO institutions that are not pilot testing or the one-eighth of PO institutions that do not know whether pilot testing is in their future are skipping directly to implementation or terminating IR-related activities is revealed by their answers to a question about their next steps pertaining to the IR effort (see Subchapter 4.4).

Figure 4.2. Likelihood of PO institutions pilot testing



4.4 Next Steps Pertaining to the IR Effort

The questionnaires asked PO respondents what steps they plan to take next as a direct result of their IR planning and asked PPT respondents what their next steps are as a direct result of their IR planning and pilot testing. Table 4.4 gives the results.

Table 4.4. Next steps pertaining to the IR effort

Next steps	PO	PPT
Your institution supports implementation of an IR software package	2	1
Your institution widens the scope of its investigation into IRs	1	2
Your institution seeks funding for the next step of investigation of IRs	3	3
Your institution seeks a partner institution(s) to share in an IR	4	4
Your institution waits for a consortium, network group, or similar to implement an IR	5	5
Your institution terminates its investigation of IRs	6	6

Ranked at or near the top for PO and PPT respondents are widening the scope of their IR investigations and implementing an IR software package, respectively. Both are logical next steps given their current stages in the IR effort.

Examining percentages of respondents' ratings conveys the strength of their convictions. Figure 4.3 graphs respondents' ratings for the top-two ranked answer categories—implementing IR software and widening the scope of planning investigations.

Two-thirds of PPT respondents at institutions said implementing IR software is "very likely" to be their next step. None says that IR software implementation is "very unlikely," and only a small percentage (2.6%) say such implementation is "somewhat unlikely." Clearly, almost all PPT respondents in the MIRACLE census will be going ahead with IR implementation.

About one-sixth of PO respondents say implementing IR software is "very likely" to be their next step. Almost 50% said it is

“somewhat likely.” Compared with PPT respondents, PO respondents are lukewarm about implementing IR software as their next step. Instead, widening the scope of their investigation into IRs is “very” (17.7%) or “somewhat” likely (54.4%) to be their next task.

Figure 4.4 graphs respondents’ ratings for the two middle-ranked answer categories—seeking funding and seeking partners. Large percentages of PO (65.8%) and PPT (55.5%) institutions will be seeking funding as their next step. Although large percentages of PO (42.3%) and even larger percentages of PPT (51.8%) respondents say they are unlikely to seek partners for IR implementation, several write-in responses mention possible participation in state-funded IRs. Not knowing their next step is more characteristic of PO respondents, about 10% of whom are not sure whether seeking funding or partners will be their next step.

Results for the bottom-ranked next steps—waiting for consortial developments and terminating IR involvement—are shown in Figure 4.5. Almost equal percentages of PO institutions are likely and are not likely to wait for a consortium or other group to implement an IR. PPT institutions appear to be speeding ahead with IR implementation—hardly 16% are waiting for a consortium or other group to implement an IR while about 70% are not waiting.

Figure 4.3. Top-ranked next steps

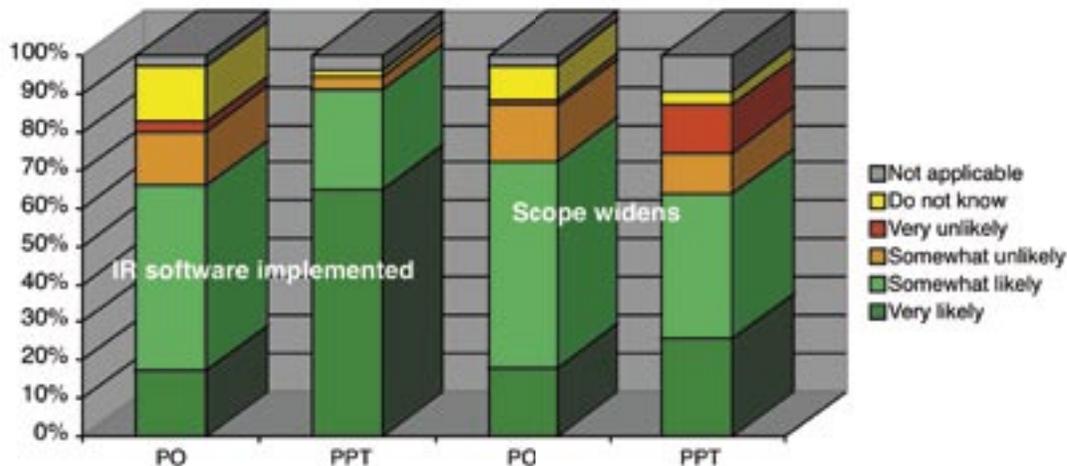
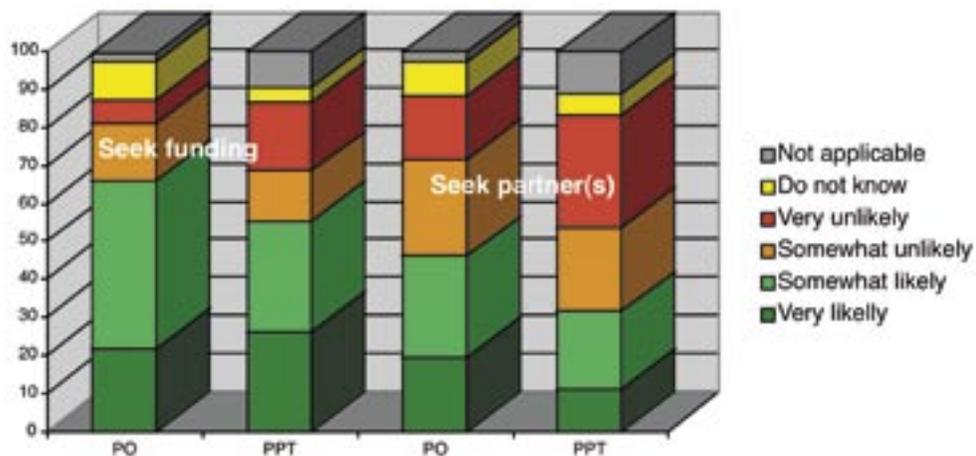


Figure 4.4. Middle-ranked next steps



Percentages of PO and PPT institutions likely to terminate their investigations of IRs are very low (10.3% and 11.2%, respectively). Some respondents misinterpreted a “very” or “somewhat” likely answer to this question to mean that they would be turning their IR investigation in a different direction, for example, toward IR pilot testing or actual IR implementation, instead of terminating all IR-related activities; consequently, the percentages of census respondents who are truly terminating may be even lower than the percentages represented in Figure 4.5. For the most part, PO and PPT respondents in the MIRACLE Project census will be continuing their institutions’ IR efforts.

Figure 4.6 shows how long it will take PO and PPT respondents to make the decision to implement an IR. Overall, PO respondents will be taking longer than PPT institutions. For example, about three-quarters of PPT respondents will be making the decision within six months. The same proportion of PO respondents will be making this same decision within 12 months.

Figure 4.5. Bottom-ranked next steps

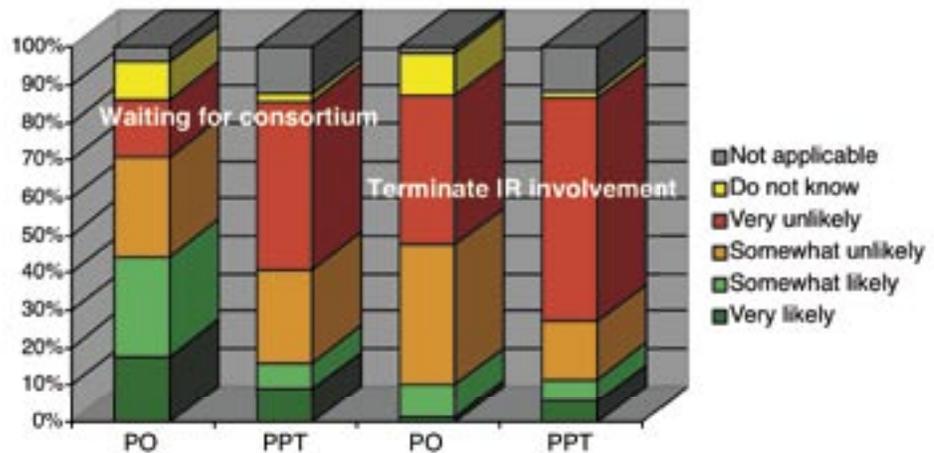
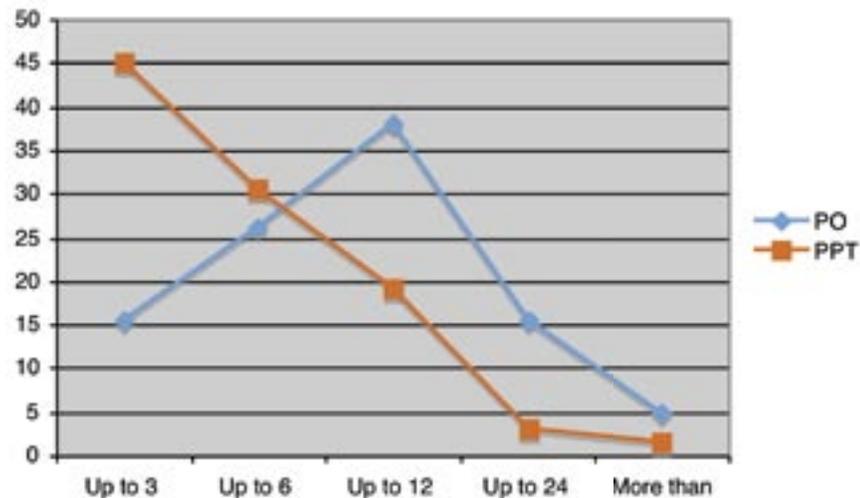


Figure 4.6. Time needed to make the IR implementation decision



4.5 Chapter 4 Summary

Asked to rate a list of 12 investigative activities, PO, PPT, and IMP respondents put those associated with learning about IRs from the experiences of others at the top. For PPT and IMP respondents, this includes analyzing literature reviews (see Table 4.1). PO respondents rank “Demonstrating operational IRs to my institution’s decision makers” much higher than PPTs and IMPs do. Such demonstrations probably make IRs more tangible to decision makers, and, possibly, more favorably inclined to support the IR effort in rhetoric and funding. In the middle of the pack is the needs assessment. Follow-up questions reveal one in sixteen PO, one in four PPT, and one in three IMP institutions, respectively, have conducted a needs assessment (see Figure 4.1). Between 5% and 12% of respondents do not know whether their institutions have conducted a needs assessment. Asked whether they are likely to conduct a needs assessment prior to making a decision about implementing an IR, about 70% of PO respondents and 44% of PPT respondents say they are “very” or “somewhat” likely to do so.

Questionnaires asked IMP respondents how important the needs assessment was for accomplishing 11 IR-related tasks (see Table 4.2). More than 75% of respondents rate all but four tasks very high in importance. At the top is “Formulating IR policies.” Because “Making the decision to implement an IR” is close to the bottom of the list, it is likely that census institutions are not conducting the needs assessment to help them make the decision to implement an IR. Instead, they are conducting it to discover how their institution’s learning community will react to the IR.

Rating the importance of various benefits of pilot testing one or more IR-system software packages, most PPT and IMP respondents choose benefits that are very practical in terms of implementing an IR—developing the requisite technical expertise, learning about IR software, and estimating costs. Middle- to low-middle ranked benefits pertain to potential contributors and users of the IR (see Table 4.3).

Census respondents in the PPT stages of the IR effort are downright positive about implementing an IR at their institutions (see Subchapter 4.4). Their next steps are widening the scope of planning activities or implementing an IR. Most will not be waiting for a consortium, partner, or group of libraries; instead, they prefer to do IR implementation on their own. Very few will be terminating all IR-related activities.

5 IR SYSTEMS AND FEATURES

Chapter 5 tells how many institutions with institutional repositories (IRs) are implementing the IR-system software packages they have chosen, describes system features that respondents believe are satisfactory and less than satisfactory, and explains why and when IMP respondents would migrate to a new IR.

5.1 Number of IRs at Institutions

On planning and pilot testing (PPT) and implementation (IMP) questionnaires, the first question asked respondents how many IRs were available or would be available to their institution's learning community in the near future. Table 5.1 lists the results.

Table 5.1. Number of IRs

Number of IRs	PPT		IMP	
	No.	%	No.	%
1	50	72.5	37	77.1
2	12	17.4	8	16.6
3	4	5.8	3	6.3
4	1	1.4	0	0.0
5 or more	2	2.9	0	0.0
Total	69	100.0	48	100.0

Most PPT and IMP institutions have one IR, but almost a quarter have two or more IRs. PPTs probably have more than one IR because they are pilot testing IR-system software packages. Also, PPTs and IMPs may be counting the academic departments and research units that have launched IR-like software to preserve, exchange, and distribute research and teaching objects among themselves, to colleagues at other schools, and to Web searchers generally. This project's phone interviews and case studies should ask follow-up questions to determine whether institutions with multiple IRs will eventually centralize IR services and, if so, which IR they will choose for centralization.

After respondents answered the first question, questionnaires instructed them to answer the remaining questions with the *one* IR in mind that offered the widest array of services to the most people and greatest number of constituencies.

5.2 IR Software

Table 5.2 enumerates the IR-system software packages that PPT and IMP respondents have pilot tested and implemented.

More institutions have pilot tested and implemented DSpace than any other IR system. This is not unexpected. DSpace was one of the first software packages specifically developed for IR services (see Appendix F2). Most Coalition for Networked Information (CNI)

members, Canadian Association of Research Libraries (CARL) members, and Association of Research Libraries (ARL)-member libraries have implemented DSpace (Lynch and Lippincott 2005; Shearer 2004; Bailey et al. 2006) (see Appendix F7).

Table 5.2. Pilot-tested and implemented IRs

System	PPT Pilot-tested IRs		IMP Pilot-tested IRs		IMP Implemented IRs	
	No.	%	No.	%	No.	%
DSpace	31	27.9	13	40.7	19	46.4
ContentDM	22	19.8	2	6.2	2	4.9
Fedora	15	13.5	3	9.4	0	0.0
Greenstone	6	5.4	3	9.4	0	0.0
Luna	6	5.4	0	0.0	0	0.0
Bepress	5	4.5	4	12.5	11	26.8
ProQuest	4	3.6	0	0.0	5	12.2
Innovative Interfaces	4	3.6	0	0.0	0	0.0
ExLibris	4	3.6	1	3.1	1	2.4
Virginia Tech ETD	3	2.7	2	6.2	1	2.4
GNU Eprints	2	1.8	1	3.1	0	0.0
Custom-made IR	2	1.8	0	0.0	2	4.9
Other	7*	6.3	3†	9.4	0	0.0
Total	111	100.0	32	100.0	41	100.0

* California Digital Library's Preservation Repository System, Confluence, Documentum, Dpubs, Endeavor, Opus Storage resources broker.

† Dpubs, Microtek Scanmaker 5, MS Access database and Cold Fusion script.

Among census respondents, Fedora and ContentDM are popular for pilot testing but not yet for implementation. Perhaps respondents have not yet had enough time to reach implementation with these packages. Developed by Berkeley Electronic Press, bepress is popular for IR implementation possibly because it hosts clients' IRs. ProQuest recently partnered with bepress to market Digital Commons, a combination of bepress and ProQuest's electronic theses and dissertations (ETDs). Academic institutions are starting to work with commercial firms such as Innovative Interfaces, ExLibris, and ProQuest in connection with their IRs.

Two IMP respondents wrote open-ended comments about replacing their in-house IRs with commercial IR products.

- "In 1998 ... only the VT-ETD system existed [and] we did not have the campus expertise to utilize [it] at that time, hence we went [with an] in-house option ... In the coming year, we will be considering commercial based IR systems."
- "[We have been] using in-house system[s] ... Presently, [we are] considering outsourcing to [a] commercial platform. Ideally, [we] would migrate to system such as Ex Libris Digital. We'll be soliciting RFPs in the coming year."

Asked how long their IR has been operational, 52.1% of IMP respondents say 12 months or less, 27.1% from 13 to 24 months, 4.2%

from 25 to 36 months, and 16.6% for more than 36 months.

IMP respondents characterize their IRs' hosts as follows: (1) their institution only, 51.2%; (2) a for-profit vendor, 31.7%; (3) a partnership that joins their institution with one or more comparable institutions, 9.8%; and (4) a regional or state-based consortium, 7.3%.

5.3 IR-system Features

Questionnaires asked PPT and IMP respondents to rate IR systems generally or their chosen IR system, respectively, with regard to various features. To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very adequate; (+1) somewhat adequate; (0) no opinion, don't know, or not applicable; (-1) somewhat inadequate; and (-2) very inadequate. They totaled the weights. These results were then compiled to rank order all the positions. Table 5.3 uses IMP ranks to order top- (1 to 5), middle- (6 to 10), and bottom-ranked (11 to 14) features.

Table 5.3. Ranking IR-system features

Top-ranked IR-system features (1 to 5)	PPT	IMP
Supported file formats	1	1
Adherence to open-access standards	2	2
Browsing, searching, and retrieving digital content	4	3
Technical support	(10)†	4
Scalability = system growth and enhancement	(7)	5
Middle-ranked IR-system features (6 to 10)	PPT	IMP
Formulating metadata for digital documents	(5)	6
Customization	8	7
User authentication	(3)	8T*
End-user interface	9	8T
Digital preservation	6	10
Bottom-ranked IR-system features (11 to 14)	PPT	IMP
Technical documentation	11	11
Extensibility = access to other campus systems and data	12	12
Controlled vocabulary searching	13	13
Authority control	14	14

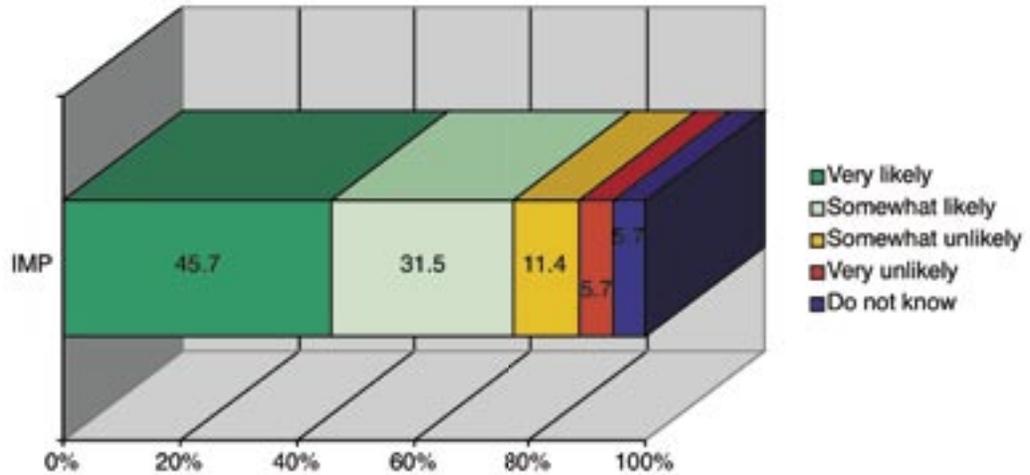
† Parentheses indicate PPT features that deviated from IMP top, middle, or bottom ranks.

* T indicates a ranked feature that tied another feature's weight.

PPT and IMP respondents agree about the two top-ranked IR-system features—supported file formats and adherence to open-access standards. PPT respondents rank technical support and the scalability of their pilot-test systems lower and rank user authentication higher than IMP respondents do. IR-system functionality for browsing, searching, and retrieving digital content is generally satisfactory; however, the user interface receives middle-ground grades. Because the user interface is usually connected to two bottom-ranked features, controlled vocabulary searching and authority control, IR systems could benefit from improvements to system features that users rely on to retrieve digital content.

Asked how likely they are to modify their IR’s software, about 75% of IMP respondents said that they are “very” or “somewhat” likely to do so (Figure 5.1). About 6% are “very unlikely” to make such modifications, and 6% do not know whether they will modify it.

Figure 5.1. Likelihood of modifying IR software



5.4 Migrating to a New IR

When asked how long they thought their institutions would stick with their present IR software before migrating to a new system, about half the respondents skipped the question. Answers from those who responded average 3.4 years. About 56% of IMP respondents think they will migrate to new IR software within the next three years. Forty percent think they will migrate in the next four to six years. The remaining 4% may stick with their present system for seven or more years.

A follow-up question asked IMP respondents to identify reasons they would migrate to a new IR. In the “% Important” column in Table 5.4 are respondents who gave a listed reason a “very” or “some-what” important rating.

Table 5.4. Reasons for migrating to a new system

Rank	Reasons	% Important
1	Greater capacity for handling preservation	90.3
2	Greater opportunities for customization	86.7
3	Greater versatility with the wide range of digital formats	80.7
4	Advanced searching features	80.7
5	Friendlier user interface	77.4
6	Better tools for assisting contributors with metadata creation	74.2
7	Friendlier digital content submission procedure	74.2
8	Greater versatility for linking to other campus systems and data	71.0
9	Around-the-clock technical support	44.0

Except for around-the-clock support, IMP respondents think every reason is important. Their top-ranked reason targets improved preservation capabilities, a feature that PPT respondents rank in the top middle of the pack and IMP respondents rank in the bottom middle of the pack with respect to their current systems (see Table 5.3). The few write-in reasons cite future availability of more commercial IR systems, opportunities to participate in an IR on the consortium level, and the cost of available software.

5.5 Chapter 5 Summary

Most institutions involved with IR PPT or implementation have one IR but almost a quarter have two or more IRs (see Table 5.1). Some PPTs may have multiple IRs because they are engaged in pilot-testing activities. In addition, respondents may be counting IR-like systems at their institutions that academic and research units have launched to share research and teaching production.

More institutions have pilot tested and implemented DSpace than any other IR system; bepress is popular for implementation (see Table 5.2). Fedora and ContentDM are popular for pilot testing but not yet for implementation. Overall, respondents have implemented more than two dozen different IR systems.

Asked how long their IR has been operational, 52.1% of IMP respondents say 12 months or less, 27.1% 13 to 24 months, 4.2% from 25 to 36 months, and 16.6% for more than 36 months.

IMP respondents characterize their IR's host as follows: (1) their institution only, 51.2%; (2) a for-profit vendor, 31.7%; (3) a partnership that joins their institution with one or more comparable institutions, 9.8%; and (4) a regional or state-based consortium, 7.3%.

PPT and IMP respondents agree on the two top-ranked IR-system features—supported file formats and adherence to open-access standards (see Table 5.3). At the bottom are controlled vocabulary searching and authority control, two features that pertain to end-user searching of IR content. Asked how likely they are to modify their IRs software, about three-quarters of IMP respondents say that they are “very” or “somewhat” likely to do so (see Figure 5.1).

IMP respondents think they will stick with their present IR system for about three-and-a-half years. Presented with a list of reasons for migrating to a new system, IMP respondents tell us that all but one (around-the-clock technical support) are important (see Table 5.4). Their top-ranked system-migration reason is greater capacity for handling preservation, a feature in their current systems that they rated in the middle (see Table 5.3).

6 IR PRACTICES AND POLICIES

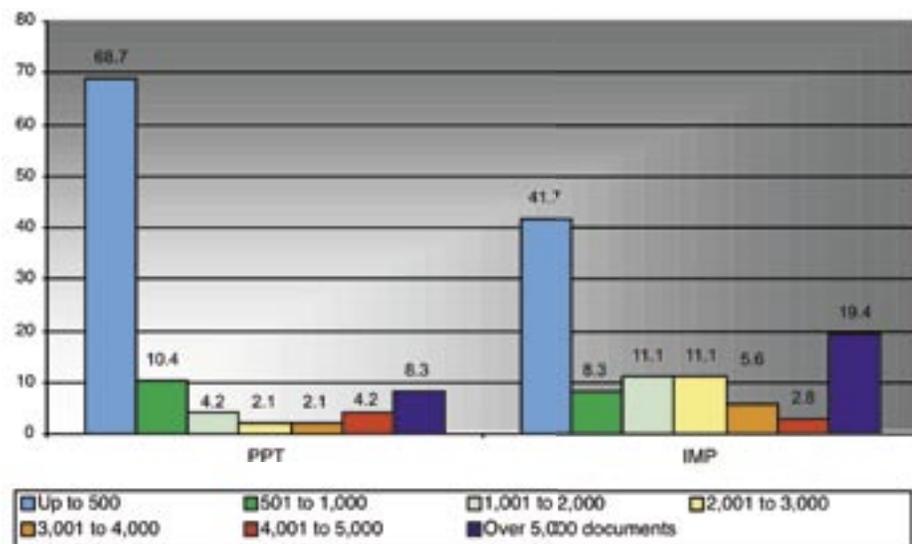
Chapter 6 gives details on institutional repository (IR) practice and policies, such as the number and types of digital documents in IRs, experiences with contributors and recruiting content for the IR, managing the IR's intellectual property (IP) rights, and the extent to which IR policies are implemented in institutions with pilot-test and operational IRs.

6.1 The Number of Digital Documents in IRs

Questionnaires asked respondents from institutions involved in IR planning and pilot testing (PPT) and IR implementation (IMP) to estimate the total number of digital documents that are published or in process in their operational or pilot-test IR. Figure 6.1 gives results.

Both pilot-test and operational IRs are very small. About 80% of PPT respondents and 50% of IMP respondents report that their IRs contain fewer than 1,000 digital documents. Only four (8.3%) IRs in the PPT stage and seven (19.4%) in the IMP stage contain more than 5,000 documents.

Figure 6.1. Number of digital documents in IRs



Thinking that older IRs would be more likely to contain more digital documents, MIRACLE Project staff compared IMP respondents' answers on this question to their responses to questions about IR age and size. Surprisingly, we did not find a relationship between IR size and age. We added PPT data to the mix and were still unable to find a relationship. On one hand, young and old IRs may have several thousand digital documents; on the other, both young and old IRs may have only a few hundred digital documents.

6.2 Digital Document Types in IRs

PPT and IMP questionnaires listed three dozen digital document types and asked respondents to estimate how many documents per type were in their respective IRs. Table 6.1 cites averages per document type and lists them in four categories from high to low. In the center “ID” column, document types in implemented IRs are numbered from 1 to 36. In the far left-hand “ID” column, these same document-type ID numbers are repeated for document types in planning and pilot-test IRs. Except for the first-listed document type, “Doctoral dissertations,” document-type ID numbers hardly ever match. Averages for some document types are surprisingly high, so we included the number of respondents who volunteered estimates. For example, “Other learning objects ... prepared by faculty, lecturers, teaching assistants, etc.” is listed third in the “PPT high” list and averages a whopping 550.0 documents per pilot-test IR; however, only four PPT respondents gave estimates and these were 0, 0, 200, and 2,000, which average to 550.0. In this case, the estimate (i.e., 2,000) given by one of a handful of respondents inflates the document type’s average.

Table 6.1. Document types in pilot-test and operational IRs

ID	PPT high: More than 200 documents	#	ID	IMP high: More than 200 documents	#
1	Doctoral dissertations (n=9)	1,288.2	1	Doctoral dissertations (n=18)	1,518.3
7	Preprints (n=7)	900.4	2	Working papers (n=18)	716.3
12	Other learning objects ... prepared by faculty, lecturers, teaching assistants, etc. (n=4)	550.0	3	Journal articles (n=19)	461.5
5	Master’s theses (n=10)	229.8	4	Raw data files that result from doctoral dissertation research (n=11)	456.6
ID	PPT medium high: 51 to 200 documents	#	5	Master’s theses (n=16)	418.8
ID	PPT medium low: 6 to 50 documents	#	ID	IMP medium high: 51 to 200 documents	#
3	Journal articles (n=14)	172.2	6	Committee meeting agenda and minutes (n=8)	90.0
2	Working papers (n=18)	124.0	7	Preprints (n=10)	84.2
30	Your institution’s course catalogs (n=7)	109.4	8	Your institution’s newspapers (n=7)	80.9
21	Books (n=4)	96.3	9	Senior theses (n=12)	68.1
20	Video recordings of performances (n=6)	76.2	10	Committee meeting documents, e.g., budgets, reports, memos (n=8)	67.5
9	Senior theses (n=7)	68.3	11	Maps (n=9)	61.1
ID	PPT medium low: 6 to 50 documents	#	ID	IMP medium low: 6 to 50 documents	#
18	Faculty senate agendas and minutes (n=5)	50.6	12	Other learning objects ... prepared by faculty, lecturers, teaching assistants, etc. (n=9)	31.0
24	Interview transcripts (n=6)	48.5	13	Written papers or transcripts of conference presentations (n=12)	27.4
33	Your institution’s alumni publications (n=4)	42.5	14	Undergraduates’ class notes, outlines, assignments, papers, and projects (n=10)	17.6
22	Sound recordings of interview transcripts (n=7)	35.9	15	Conference presentations (e.g., summaries, abstracts, notes, outlines) (n=10)	16.1
8	Your institution’s newspapers (n=2)	33.8			

27	Course syllabi, class notes, handouts, etc., prepared by faculty, lecturers, etc. (n=3)	33.3	16	Interim and final reports to funders (n=9)	12.8
19	Regent, trustee, board meeting agenda and minutes (n=4)	30.0	17	College, departmental, and school alumni publications (n=9)	12.6
10	Committee meeting documents, e.g., budgets, reports, memos (n=5)	28.2	18	Faculty senate agendas and minutes (n=8)	12.5
6	Committee meeting agenda and minutes (n=5)	27.8	19	Regent, trustee, board meeting agenda and minutes (n=9)	11.1
17	College, departmental, and school alumni publications (n=4)	22.3	20	Video recordings of performances (n=12)	9.4
23	Journals (n=16)	22.0	21	Books (n=15)	6.3
13	Written papers or transcripts of conference presentations (n=3)	17.0	22	Sound recordings of interview transcripts (n=11)	6.0
15	Conference presentations (e.g., summaries, abstracts, notes, outlines) (n=4)	7.8	ID	IMP low: 5 or fewer documents	#
34	Undergraduate student e-portfolios (n=2)	7.7	23	Journals (n=6)	4.0
ID	PPT low: 5 or fewer documents	#	24	Interview transcripts (n=10)	3.3
11	Maps (n=3)	5.0	25	Raw data files that result from master's thesis research (n=9)	2.3
31	Raw data files from senior thesis research (n=3)	5.0	26	Software (n=9)	2.2
29	Raw data files from faculty research projects (n=5)	3.3	27	Course syllabi, class notes, handouts, etc., prepared by faculty, lecturers, etc. (n=9)	1.8
26	Software (n=2)	2.5	28	Software documentation (n=9)	1.7
14	Undergraduates' class notes, outlines, assignments, papers, and projects (n=2)	0.0	29	Raw data files from faculty research projects (n=7)	1.4
4	Raw data files that result from doctoral dissertation research (n=1)	0.0	30	Your institution's course catalogs (n=7)	1.4
16	Interim and final reports to funders (n=1)	0.0	31	Raw data files from senior thesis research (n=8)	1.3
25	Raw data files that result from master's thesis research (n=1)	0.0	32	Graduate students' class notes, outlines, assignments, papers, and projects (n=8)	0.8
28	Software documentation (n=1)	0.0	33	Your institution's alumni publications (n=9)	0.0
32	Graduate students' class notes, outlines, assignments, papers, and projects (n=1)	0.0	34	Undergraduate student e-portfolios (n=2)	0.0
35	Graduate student e-portfolios (n=2)	0.0	35	Graduate student e-portfolios (n=7)	0.0
36	Blogs (n=1)	0.0	36	Blogs (n=8)	0.0
PPT average and total (n=4.9)		4,038.0	IMP average and total (n=9.9)		4,206.4

Although MIRACLE Project investigators were skeptical about including such a long list of digital types in the questionnaires, we are glad we did because the results show a wide range of document types in both pilot-test and operational IRs. Estimates of the various document types in both pilot-test and operational IRs are generally low, seldom exceeding 50 documents per type. Estimates for pilot-test and operational IRs are not that much different; in fact, totaling the two estimates results in hardly a 200-document difference in favor of operational IRs. With a few exceptions (e.g., the "Other learning objects..." type described above), greater numbers of respon-

dents volunteering estimates result in document types listed in Table 6.1's "high" and "medium-high" categories and lower numbers of respondents volunteering estimates result in document types listed in Table 6.1's "medium-low" and "low" categories.

At the top of both the PPT and IMP lists are the traditional text-based document types that are the result of the research enterprise of faculty and students at postsecondary institutions, e.g., doctoral dissertations, master's theses, working papers, preprints, and journal articles. Large numbers of doctoral dissertations, raw data files that result from doctoral dissertation research, and master's theses in the IR may be the result of institutional monitoring of student compliance with mandatory submission of these document types. Respondents do not always give high estimates for document types that would be packaged in numeric and multimedia files (e.g., video recordings of performances, e-portfolios, raw data files, software, sound recordings of interview transcripts, maps), but there is evidence that numbers for nontext files will grow in the years to come.

Census respondents volunteered document types we missed. These included government documents, archives, institutional historical documents (including photographs and art history slide collections), faculty spatial data sets, staff project reports, research reports from centers and institutes, self-study reports, and other documentation from academic accreditation events, posters, newsletters, musical scores, and scrapbooks.

6.3 Status of IR Policies

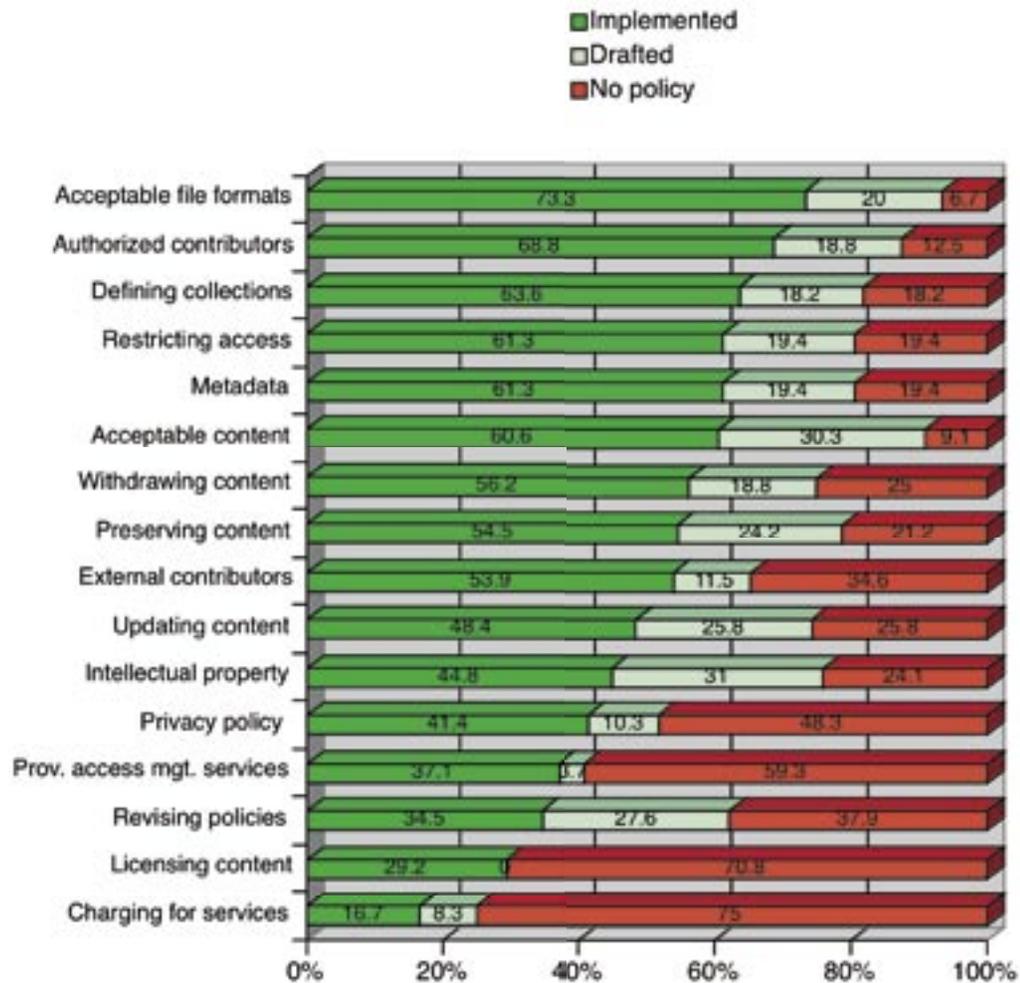
PO, PPT, and IMP questionnaires asked respondents to characterize the status of 16 policies as follows: (1) no policy, (2) drafted policy, (3) implemented policy, (4) do not know, and (5) not applicable.

Sixteen IMP respondents skipped the policy question and three chose the "do not know" or "not applicable" categories. As a result, about 60% of the total 48 IMP respondents answered this question. Figure 6.2 shows the status of policies at IMP institutions.

High percentages of IMP respondents report implemented policies for (1) acceptable file formats (73.3%), (2) determining who is authorized to make contributions to the IR (68.8%), (3) defining collections (63.6%), (4) restricting access to IR content (61.3%), (5) identifying metadata formats and authorized metadata creators (61.3%), and (6) determining what is acceptable content (60.6%). Generally, IMP institutions have implemented or drafted policies for all but these four policies: (1) charging for IR services (16.7%), (2) licensing IR content (29.2%), (3) formulating a privacy policy for registered IR-system users (41.4%), and (4) providing access management services (37.1%).

It may be not necessary for all IR policies to be in place for IR implementation. IMP institutions may take a wait-and-see attitude, evaluating what transpires after a period of time, and then firming up existing policies and implementing new ones as needed. Whether this is the case will be verified in the phone interviews and case studies to be conducted in later phases of the MIRACLE Project.

Figure 6.2. Status of IR policies at IMP institutions

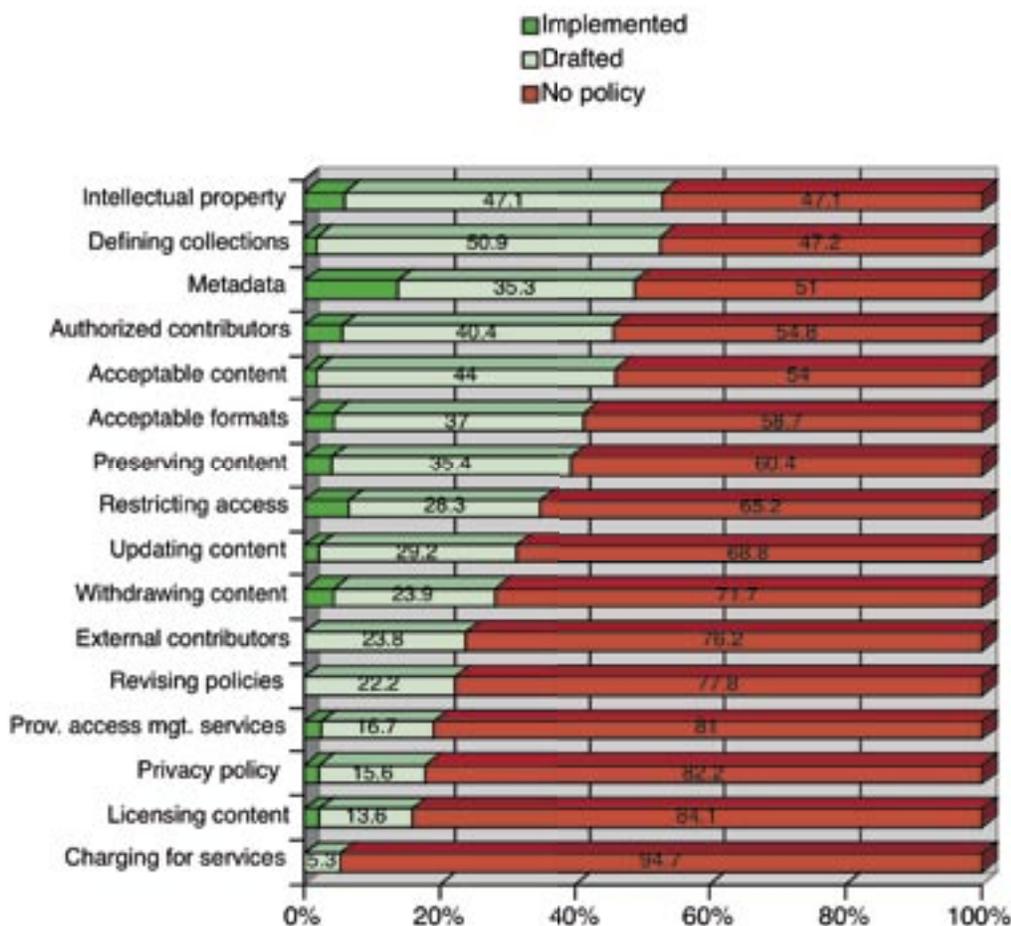


Sixteen PPT respondents skipped the policy question altogether and another nine chose the “do not know” or “not applicable” categories. As a result, about 64% of the total 70 PPT respondents answered this question. Figure 6.3 shows the status of policies for PPT institutions.

Except for metadata formats, percentages of *implemented* policies are in the single digits at PPT institutions. Between one-third and one-half of such institutions have, however, *drafted* policies in the following six areas:

1. Defining collections (1.9% implemented, 50.9% drafted)
2. Intellectual property (5.9% implemented, 47.1% drafted)
3. Determining what is acceptable content (2.0% implemented, 44.0% drafted)
4. Determining who is authorized to make contributions to the IR (5.8% implemented, 40.4% drafted)
5. Acceptable file formats (4.4% implemented, 37.0% drafted)
6. Metadata formats and authorized metadata creators (13.7% implemented, 35.3% drafted)

Figure 6.3. Status of IR policies at PPT institutions



Between 60% and 95% of PPT institutions have no policies whatsoever for the bottom 10 policies listed in Figure 6.3. A handful of respondents who replied to the open-ended component of this question say as much:

- “Nothing implemented yet—still investigating.”
- “Too soon to tell.”
- “We are just in the planning stage now. We are just starting.”
- “No policy yet written.”

A few PO respondents who have gotten started on policy formulation report that they have drafted policies. One or two have implemented policies. They started with the same policies as PPT respondents did, namely:

1. Intellectual property (2.8% implemented, 12.5% drafted, 84.7% no policy)
2. Determining what is acceptable content (1.4% implemented, 11.3% drafted, 87.3% no policy)
3. Acceptable file formats (1.4% drafted, 88.6% no policy)
4. Defining collections (11.4% drafted, 88.6% no policy)
5. Determining who is authorized to make contributions to the IR (11.1% drafted, 88.9% no policy)
6. Identifying metadata formats (9.9% drafted, 90.1% no policy)

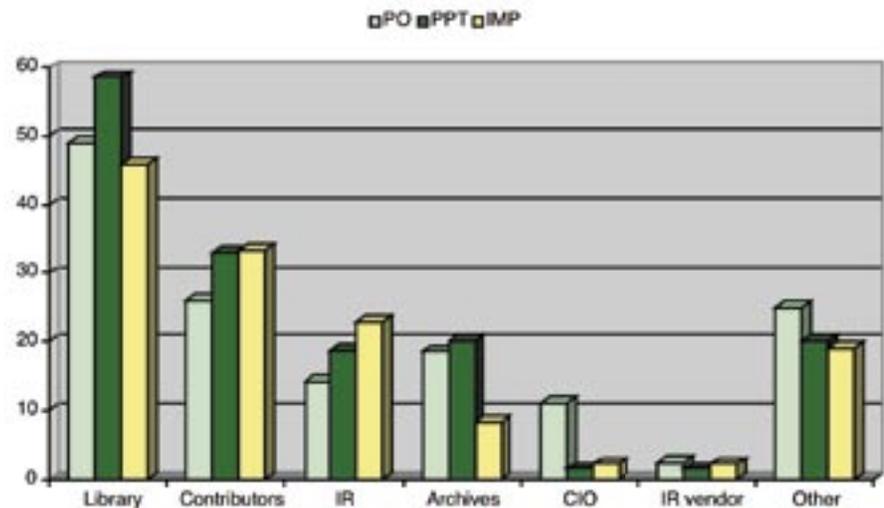
The percentage of respondents who report no policy for the remaining 10 policies ranges from 93% to 100%. Respondents at PO institutions say that it is too early to get started on policy formulation.

- “Given the stage at which the institution is at [with regard to IRs], no policies are in place.”
- “We are truly at the very beginning stages of examining this issue. We are sending a group of librarians/faculty to a scholarly communication workshop next month and will use that core group to begin investigating what we will do here.”
- “We are not yet far enough along in the planning process to be able to effectively and prudently draft policies.”

PO, PPT, and IMP questionnaires asked respondents who manages the IR’s IP rights. Because respondents could choose more than one response category, responses exceed 100%. Figure 6.4 gives the results. It reveals that library staff manage the IR’s property rights. IP rights are also in the hands of the contributors themselves. Here is what respondents had to say in this regard:

- “We basically ask our contributors to sign a license agreement that says they’ve cleared any copyright issues. So, managing the IR’s intellectual property rights is not a huge task.”
- “Authors posting their materials [in the IR] should clarify copyright status.”
- “Contributors decide.”
- “Contributor, especially regarding copyright.”
- “Submitters.”

Figure 6.4. Managing IP rights in the IR



PPT and IMP respondents volunteer “Others,” such as the university counsel and the institution’s IR vendor. Most PO and about half of PPT respondents who checked “Other” confess that they do not know, are unsure, or are not far enough along in the planning process to know the answer.

Managing IP rights can be a partnership involving IR staff who

profile their institution's IR, the system's IP process, and the contributor who encounters this process when depositing content into the IR. Respondents' uncertainty about IP rights may be a result of their lack of familiarity with how their IR systems usher contributors through the IP process.

The domain of IP rights in IRs deserves more coverage than can be done through Web-administered questionnaires. MIRACLE Project staff will learn more about IP in their subsequent activities.

6.4 File Formats Guaranteed in Perpetuity

PPT and IMP questionnaires listed 34 file formats and asked respondents to check which ones their IR guarantees in perpetuity. Few PPT respondents were prepared to be definitive about their responses, checking instead answer categories such as "do not know," "no opinion," or "not applicable." In write-in responses, they said they could make guarantees at this time:

- "We don't guarantee anything, at least not until we have a preservation plan in production."
- "No guarantees yet. Too soon to say."
- "I don't know. ProQuest will handle the technical part."
- "We will preserve many of these formats, but not in perpetuity. Storage options are changing too rapidly."

Table 6.2. Guaranteed digital file formats

File format	Skipped	*Other	Guaranteed by IMP respondents	
	No.	No.	No.	%
PDF	15	6	27	56.3
JPEG	24	6	18	37.5
TIFF	22	8	18	37.5
GIF	24	10	14	29.2
XML	24	11	13	27.1
Microsoft Word	25	10	13	27.1
Microsoft Excel	26	11	11	22.9
PDF/A	22	15	11	22.9
Rich text	26	12	10	20.8
Microsoft PowerPoint	26	12	10	20.8
Postscript	24	15	9	18.8
MPEG audio	26	14	8	16.7
Plain text ANSI X3.4/ECMA-6/US-ASCII (7-bit)	26	15	7	14.8
Plain text UTF-8 (Unicode)	25	16	7	14.8
Plain text ISO 8859-x (8-bit)	26	16	6	12.5
Plain text (all other encodings, including, but not limited to, ISO 646, national variants)	26	16	6	12.5
PNG	25	17	6	12.5
TeX	25	17	6	12.5

* Other includes responses for "do not know," "no opinion," and "not applicable."

Because PPT respondents are not yet far enough along in the planning process to discuss guaranteed file formats with a high degree of certainty, we excluded them from Table 6.2. This table enumerates digital file formats that at least 12.5% of IMP respondents guarantee in perpetuity. Because IMP respondents also appear to be uncertain about guaranteeing file formats, Table 6.2 includes the number of IMP respondents who failed to answer the question.

Except for PDF files, about half of IMP respondents skipped the question. Most Table 6.2 file formats handle text, numerical, or image data. Only one handles audio data. Generally fewer than 12.5% of IMP respondents guarantee multimedia formats such as QuickTime (10.4%); MPEG-4 (10.4%); Windows Media Video (6.3%) and AVI (6.3%); sound formats such as AIFF (8.3%), Real Audio (6.3%), and Wave (6.3%); and image formats such as BMP (10.4%) and PhotoCD (6.3%).

One respondent comments specifically about MIRACLE census questions pertaining to preservation and about digital preservation generally:

- “It ... depends on what you mean by ‘preserved.’ We will be providing bit-level preservation for all [listed] formats, but aren’t promising that the files will be usable in terms of software available. There needs to be a better definition of digital preservation in this question—it’s not just a matter of supported and unsupported formats—we have a rather complicated system for determining levels of support. I’m a little put off by the questions about digital preservation in this survey since no IR software that I know of (nor any digital content management software for that matter) provides off-the-shelf digital preservation capabilities. It takes a lot of work to build the additional infrastructure and identify the additional resources and policies needed to actually preserve items.”

MIRACLE Project investigators wanted to question census respondents about digital preservation issues. Although questionnaire drafts contained a number of long and complicated questions about preservation, only one question on this subject survived the editing and review process. Although simple, it revealed the large measure of uncertainty about preservation. Here is what we learned from it:

- Few PPT respondents are prepared to address long-term preservation issues.
- Except for PDF files, percentages of IMP respondents guaranteeing file formats in perpetuity are low.
- Percentages of IMP respondents guaranteeing in perpetuity some image formats and almost all audio and video file formats are very low (i.e., less than 12.5%).

Having learned from census results, MIRACLE Project investigators will make digital preservation a major theme in subsequent activities.

6.5 Contributors to IRs

6.5.1 Authorized Contributors

All questionnaires asked respondents about authorized contributors to IRs, but they asked them in different ways:

- NP: If your institution eventually does make the decision to implement an IR, who do you think would be authorized contributors?
- PO and PPT: If you could foretell the future, who will be authorized contributors to your institution's IR?
- Who are authorized contributors to your institution's IR?

Questionnaires listed a dozen choices, and respondents checked ones that were or were likely to be authorized IR contributors. Table 6.3 gives the results. It uses IMP percentages to list contributors in order from high to low, and its four "Rank" columns to the left of the "%" columns rank percentages for each of the four listed respondent types.

Table 6.3. Authorized contributors to IRs

Contributor	NP		PO		PPT		IMP	
	Rank	%	Rank	%	Rank	%	Rank	%
Librarians	2	82.2	2	85.9	2	85.7	1	79.2
Faculty	1	89.0	1	87.0	1	91.4	2	77.1
Graduate students	5	45.8	5	63.0	4T	64.3	3T	56.3
Research scientists	11	26.7	7T	46.7	7	57.1	3T	56.3
Archivists	4	61.4	3	80.4	3	78.6	5	54.2
Undergraduate students	6	40.7	6	51.1	6	58.6	6	43.8
Your institution's administrators	3	65.3	4	65.2	4T	64.3	7	39.6
Academic support staff	9	34.7	7T	46.7	9	50.0	8	37.5
Your institution's central computer services staff	8	37.3	10	40.2	11	31.4	9T	16.7
Your institution's press	10	27.5	11	31.5	10	44.3	9T	16.7
External contributors	12	11.9	12	14.1	12	21.4	9T	16.7
Your institution's news service	7	38.1	9	45.7	8	52.9	12	12.5

Faculty and librarians top the list for all four respondent types. Librarians and archivists are especially likely to be active contributors on their own because they have work assignments connected with digitizing and depositing special collections in the IR. In addition, they may be proxies for faculty and research scientists who want to deposit content in the IR but have no time to do it. Surprisingly, at IMP institutions, archivists fall in the middle of the pack—below research scientists and graduate students. Why the archivist is a middle-of-the-pack authorized contributor is something that MIRACLE Project staff will pursue in subsequent activities.

Large percentages of IMP respondents acknowledge only "Research scientists" as contributors. Middle-of-the-pack percentages come from PO and PPT respondents, and NP respondents put re-

search scientists well toward the bottom of the list. Research scientists typically staff research institutes, centers, and clinical units at research-intensive universities. Such scientists would be especially prolific at generating data sets and writing reports, white papers, conference presentations, and journal articles that would be appropriate for deposit in IRs. A large percentage of participating IMP institutions are classed as research universities (see Table 2.3); for this reason, many IMP respondents authorize research scientists as IR contributors. Research scientists are less likely to be members of the learning communities at baccalaureate and master's institutions, and such institutions are more typical of NP, PO, and PPT institutions in the MIRACLE Project census.

The percentages of NP, PO, and PPT respondents authorizing college and university administrators as IR contributors are somewhat higher than the percentages of IMP respondents. Maybe the former are "playing up" to administrators because they need their support to implement an IR. It could also be that the administrators who serve on IR planning committees express greater optimism about the potential of IR contributions from fellow administrators than is actually present.

Less likely to be authorized as IR contributors are the institution's news service, press, central computer services staff, academic support staff, and external contributors. External contributors rank at or almost at the bottom. A few write-in responses from IMP respondents mention that anyone—even external contributors—could submit material to their IRs as long as a faculty member or academic department is willing to sponsor them. Several write-in responses from NP, PO, and PPT respondents mention that alumni may be authorized contributors to their IRs.

6.5.2 *The Major Contributor to the IR*

Questionnaires asked respondents who they thought would be (PO and PPT) or who is (PPT) *the* major contributor to their IR. Respondents could choose only one answer category. Table 6.4 gives the results.

Table 6.4. *The major contributor to the IR*

Major contributor	PO		PPT		IMP	
	No.	%	No.	%	No.	%
Faculty	39	48.1	37	59.7	13	33.3
Graduate students	4	4.9	2	3.2	8	20.5
Librarians	9	11.1	12	19.4	4	10.3
Undergraduate students	7	8.6	3	4.8	3	7.7
Research scientists	2	2.5	2	3.2	3	7.7
Archivists	16	19.8	5	8.1	3	7.7
Academic support staff	0	0.0	0	0.0	2	5.1
Central administrators	2	2.5	1	1.6	1	2.6
Other	2	2.5	0	0.0	2	5.1
Total	81	100.0	62	100.0	39	100.0

Although IMP respondents credit faculty with being the major contributors to their IRs, they are not overly optimistic about faculty contributions. In fact, only 33.3% of IMP respondents choose faculty as the major contributors to their IRs. PO and PPT respondents are much more positive about faculty contributions, with percentages coming close to 50% and 60%, respectively.

PO and PPT respondents do not foresee graduate students being major contributors to their IRs, but graduate students are major contributors at IRs at some IMP institutions. A large percentage of PO respondents envision archivists being major IR contributors; however, IMP respondents do not perceive archivists to be as active as other contributors to their operational IR. The large percentage of PPT respondents who choose librarians may be a result of the added workload librarians assume during the planning and PPT phase—publicizing the IR, identifying first adopters by submitting content to the IR on behalf of faculty and students, and engaging in similar activities. Only 10.3% of IMP respondents choose librarians as their IR's major contributor. Into the "Other" category a couple of IMP respondents write about the contributions of "Publishers" and a unique collection contributed by a local association and a couple of PO respondents write about the "Media relations department" and "Academic support staff."

6.5.3 Early Adopters of IR Technology

Asked about early adopters of IR technology, about two-thirds of PPT and one-half of IMP respondents have worked with their institution's library or a particular academic college, department, or school. Examples are academic units in the humanities, laboratories, centers, an undergraduate symposium, and the graduate school with regard to dissertations and master's theses. About two-fifths of PPT and one-quarter of IMP respondents have worked with their institutions' archives.

Questionnaires asked PO and PPT respondents what digital content-recruitment methods they thought would be most successful at their institution, and asked IMP respondents to assess their methods of recruiting digital content for the IR. The majority of PO, PPT, and IMP respondents gave "very successful" ratings to only one of the nine listed methods—"Staff responsible for the IR working one-on-one with early adopters." Figure 6.5 shows respondents' ratings in this regard. IMP respondents were less positive about this method, with 61.1% giving it a "very successful" rating compared with PO and PPT respondents (72.5% and 79%, respectively).

Another successful content-recruiting method is word of mouth from early adopters to their colleagues in the faculty and staff ranks (Figure 6.6). IMP respondents who check the "do not know" or "no opinion" categories may find it difficult to assess this method's success because, unlike the other recruitment methods above and below, this method does not involve them personally. PO and PPT respondents are more positive about the word of mouth method, but they probably have less recruiting experience than IMP respondents do.

Figure 6.5. Recruiting content by working with early adopters

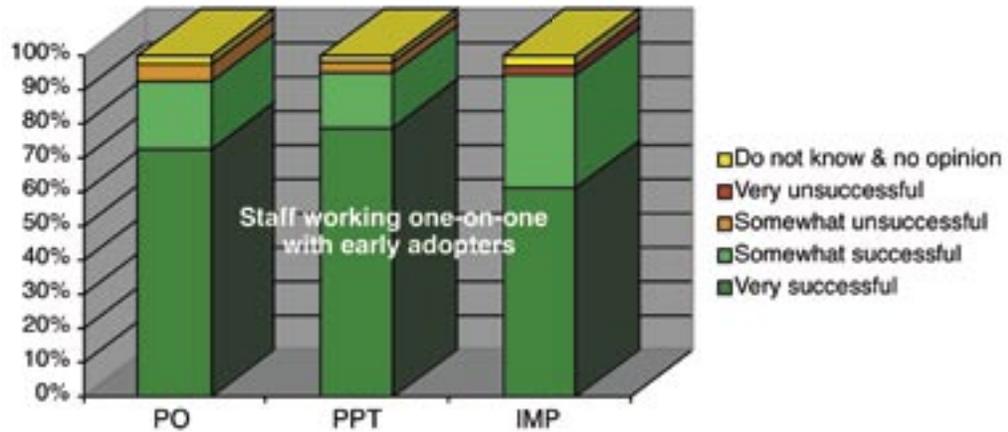


Figure 6.6. Recruiting content by word of mouth

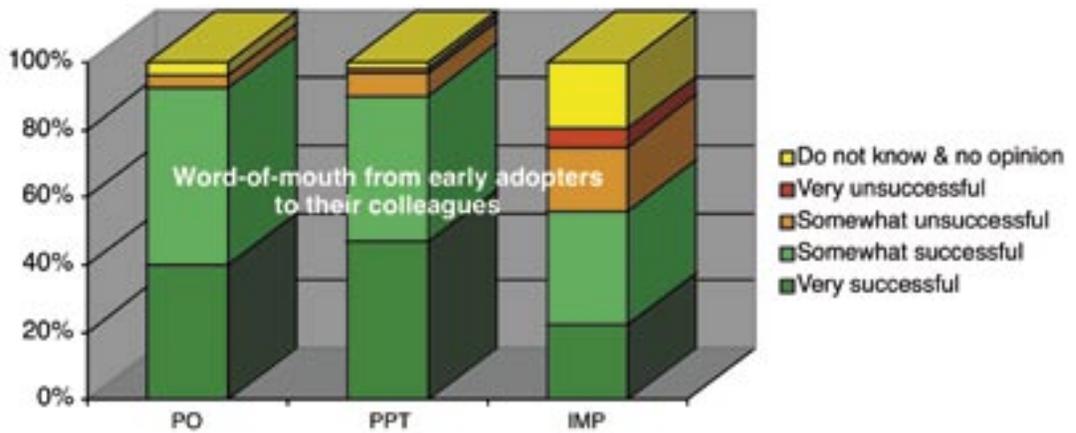
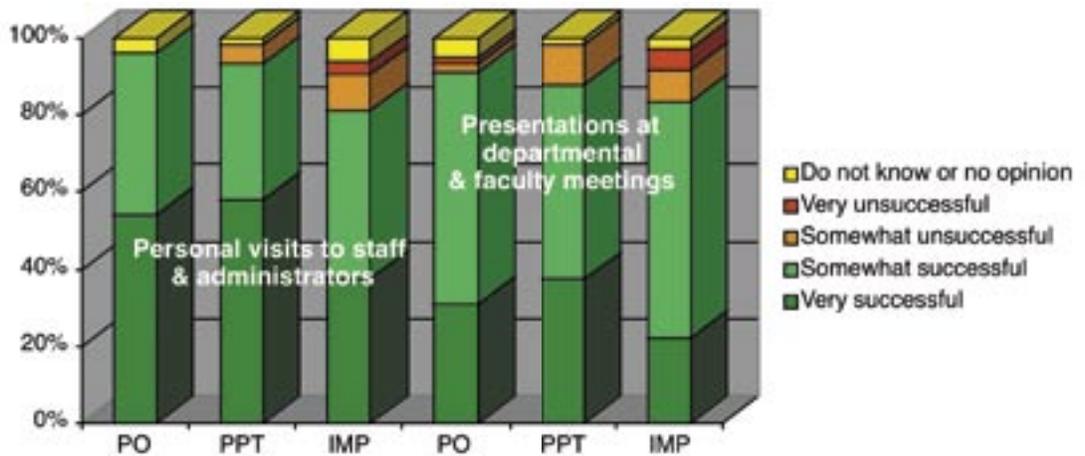


Figure 6.7. Recruiting content by personal visits and presentations



6.5.4 Other Digital Content Recruitment Methods

Ratings respondents give to other digital content recruitment methods are also of interest. Making personal visits to faculty and administrators and giving presentations about the IR at departmental and faculty meetings are less successful than working one on one with early adopters; however, respondents are still quite positive about their success using these two methods, especially the former (Figure 6.7). Perhaps the small but measurable percentage of IMP respondents checking “do not know” or “no opinion” did not engage in personal visits or presentations.

Questionnaires asked respondents about five other content-recruitment methods, listed below. Most IMP respondents rated them “successful”; however, percentages of “unsuccessful” ratings are sometimes sizable, ranging from 20% to 34%. The only method of content recruitment for which most PO and PPT respondents were negative was publicizing the IR in campus newspapers.

- volunteer contributions
- institution-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master’s theses, faculty preprints
- systematic review of faculty, staff, center, and departmental Web sites for potential contributors by staff responsible for the IR publicizing the IR during reference interactions in libraries and archives
- publicity about the IR in campus newspapers

In write-in comments, respondents volunteer a few methods that we failed to include on the questionnaires:

- “Referral from departments doing somewhat related work.”
- “Open house targeted at faculty with content.”

Write-in comments from three respondents tell how they are harvesting publicly available materials for their IRs:

- “Recently, the dean of the library approved harvesting of publicly available materials by the IR staff.”
- “Review of other open-access sites for potential IR content to be added by library staff.”
- “Prepopulation of the repository with materials from PubMed and other open databases that allow IRs to download content.”

One IMP respondent is adamant about the failure of the voluntary method:

- “So far we have made this a voluntary effort for faculty and for undergraduates. It has not really caught on for faculty and it has not been made mandatory for undergraduate theses. Voluntary does not work.”

6.6 Chapter 6 Summary

Both pilot-test and operational IRs are very small, but they contain a wide range of digital document types—text, image, audio, video, and data files (see Figure 6.1 and Table 6.1). About 80% of the pilot-test and 50% of operational IRs contain fewer than 1,000 digital documents. In the MIRACLE Project census are four (8.3%) pilot-test IRs and seven (19.4%) operational IRs containing more than 5,000 documents. There is no relationship between IR size and age. Young and old IRs may have several thousand digital documents or only a few hundred.

IRs in both the pilot-test and operational stages bear the traditional text-based document types that result from the research enterprise of faculty and students at postsecondary institutions, e.g., doctoral dissertations, master's theses, working papers, preprints, and journal articles (see Table 6.1). Estimates of the various document types in both pilot-test and operational IRs are generally low, seldom exceeding 50 documents per type. Adding up average estimates for PPT and IMP IRs reveals that an average IR bears about 4,100 digital documents representing about 30 document types. The difference between pilot-test and operational IRs may be as small as 200 documents.

Respondents from institutions where IRs have been implemented have made the most progress on IR policy (see Figure 6.2). More than 60% of IMP respondents report implemented policies for (1) acceptable file formats (73.3%), (2) determining who is authorized to make contributions to the IR (68.8%), (3) defining collections (63.6%), (4) restricting access to IR content (61.3%), (5) identifying metadata formats and authorized metadata creators (61.3%), and (6) determining what is acceptable content (60.6%). Large percentages of IMP institutions have implemented or drafted policies for all but these four policies: (1) charging for IR services, (2) licensing IR content, (3) providing access-management services, and (4) formulating a privacy policy for registered IR system users.

At PPT institutions, the emphasis is on drafting, not implementing, policy (see Figure 6.3). Between a third and a half of PPT respondents report they have drafted policies pertaining to collections, IP, acceptable content, metadata, acceptable file formats, and authorized contributors. A few PO respondents report having drafted policies. PO respondents have started with the same policies as those on which PPT respondents report the most progress.

Asked who manages the IR's IP rights, about 50% of PO, PPT, and IMP respondents cite library staff and about 30% cite the contributors themselves (see Figure 6.4).

Few PPT respondents are prepared to guarantee specific file formats in perpetuity (see Table 6.2). The only file formats for which there is much certainty about long-term guarantees are PDF, JPEG, TIFF, and GIF files in operational IRs.

At PO, PPT, and IMP institutions, authorized contributors to the IR are faculty, librarians, graduate students, research scientists, and archivists (see Table 6.3). Less likely to be authorized as IR contributors are the institution's news service, press, central computer services staff, academic support staff, and external contributors.

Although IMP respondents credit faculty with being *the* major contributor to the IR, they are not overly optimistic about faculty contributions, as reflected in the fact that only 33.3% of IMP respondents choose faculty as the major contributor to their IRs (see Table 6.4). PO and PPT respondents are much more positive about faculty contributions, with percentages coming close to 50% and 60%, respectively.

Asked about early adopters of IR technology, about two-thirds of PPT and half of IMP respondents have worked with their institution's library or a particular academic college, department, or school. About two-fifths of PPT and one-quarter of IMP respondents have worked with their institution's archives.

The most successful digital content-recruitment method is staff working one-on-one with early adopters (see Figure 6.5). Other successful methods are word of mouth from early adopters to their colleagues in the faculty and staff ranks (see Figure 6.6), personal visits by IR staff to faculty and administrators, and presentations by IR staff at departmental and faculty meetings (see Figure 6.7).

7 IR BENEFITS AND BENEFICIARIES

Chapter 7 explores institutional repository (IR) benefits from the perspectives of IR staff as well as staff perceptions of IR contributors. It also examines the IR's effect on building relationships with other units, the deployment of a successful IR, and the methods that institutions are using to evaluate IRs.

7.1 Benefits of IRs

Questionnaires asked all census respondents about the benefits of IRs but they asked the question in different ways depending on the extent of respondents' involvement with IRs:

- No planning (NP): How important do you think these anticipated benefits of IRs would be to your institution?
- Planning only (PO) and planning and pilot testing (PPT): How important are these anticipated benefits of IRs to your institution?
- Implementation (IMP): At the beginning of IR planning at your institution, how important did you think these anticipated benefits of IRs would be to your institution?

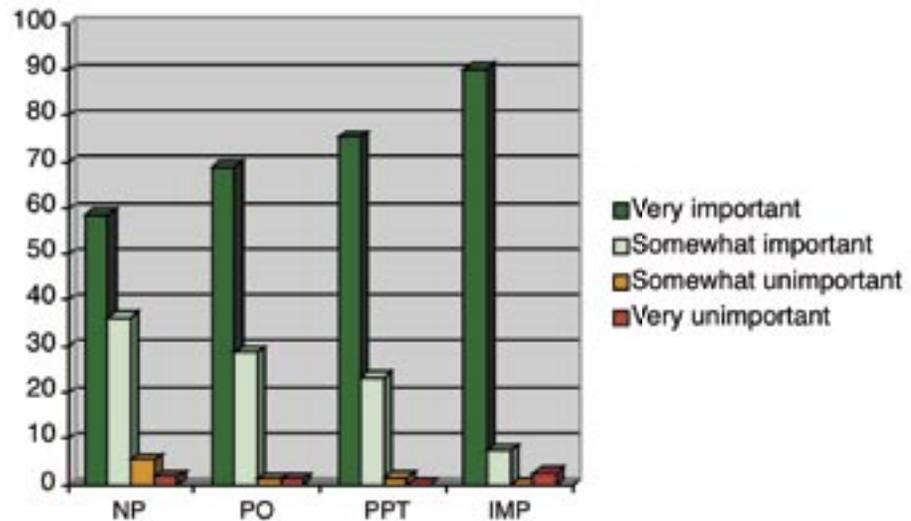
All four questionnaires listed the same 16 anticipated benefits and same response categories. With two exceptions, respondents were uniformly positive about rating listed benefits. When the percentages of "very" and "somewhat" important ratings were totaled, the sum equaled or exceeded 67% for 14 of the 16 benefits.

Respondents ranked the following two benefits next to last and last, respectively:

- an increase in citation counts to your institution's intellectual output
- reducing user dependence on your library's print collection

Respondents' positive ratings vary in a systematic way. IMP respondents' ratings for a listed benefit are almost always greater than PPT respondents' ratings for a listed benefit. Likewise, PPT respondents' ratings for a listed benefit are almost always more positive than PO respondents' ratings. Finally, PO respondents' ratings for a listed benefit almost always exceed NP respondents' ratings. Even though NP respondents are not as positive as respondents involved with IRs (i.e., POs, PPTs, and IMPs), they are definitely positive about the ratings they give to IR benefits. Figure 7.1 shows how respondents' positive ratings for listed benefits increases from NP to PO, from PO to PPT, and, finally, from PPT to IMP.

Figure 7.1. Increasingly positive nature of respondents' ratings for IR benefits



These data beg the question—why should IMP respondents be more positive about IR benefits than respondents in the other groups? It may be that IMP respondents, having experienced the IR implementation effort from beginning to end, are more confident about IR benefits and that they express this confidence by giving benefits high ratings. Or, having invested much time and effort into IR implementation, IMP respondents may want the IR to succeed so strongly that they give it the highest ratings.

To simplify results of this analysis of benefits, MIRACLE Project staff assigned weights to response categories as follows: (+2) very important; (+1) somewhat important; (0) no opinion, don't know, or not applicable; (-1) somewhat unimportant; and (-2) very unimportant. They added up the weights. These results were then compiled to rank order all the positions. Table 7.1 uses IMP ranks to order top- (1 to 7), middle- (8 to 14), and bottom-ranked (15 to 16) benefits.

Table 7.1. IR benefits

Top-ranked benefits (1 to 7)	NP	PO	PPT	IMP
Capturing the intellectual capital of your institution	2	2	2	1
Better service to contributors	(8)†	6	3	2
Exposing your institution's intellectual output to researchers around the world who would not have access to it through traditional channels	(9)	(9)	(7)	3
An increase in your library's role as a viable partner in the research enterprise	6	5	4	4
Longtime preservation of your institution's digital output	3	3	5	5T*
Better services to your institution's learning community	1	1	1	5T
A solution to the problem of preserving your institution's intellectual output	5	4	6	7
Middle-ranked benefits (8 to 14)	NP	PO	PPT	IMP
An increase in the accessibility to knowledge assets such as numeric, video, audio, and multimedia formats	(7)	8	8	8
A boost to your institution's prestige	14	13	10	9
Maintaining control over your institution's intellectual property	(4)	(7)	9	10
Contributing to the reform of the entire enterprise of scholarly communication and publishing	13	14	12	11
New services to learning communities beyond your institution	10T	10	11	12
A reduction in the amount of time between discovery and dissemination of research findings	12	11T	13	13
Providing maximal access to the results of publicly funded research	10T	11T	14	14
Bottom-ranked benefits (15 to 16)	NP	PO	PPT	IMP
An increase in citation counts to your institution's intellectual output	15	15	15	15
Reducing user dependence on your library's print collection	16	16	16	16

† Parentheses indicate NP, PO, and PPT benefits that deviated from IMP top, middle, or bottom ranks.

* T indicates a ranked benefit that tied another benefit's weight.

IMPs do not agree with NPs, POs, and PPTs about the top-ranked benefit. IMPs choose "Capturing the intellectual capital of your institution" and three others choose "Better services to your institution's learning community." However, all four respondent types are very positive about both benefits.

Questionnaires asked IMP respondents about benefits a second time. The question was "Now that you are implementing or have implemented an IR, reassess these same anticipated benefits of IRs and tell whether you think they are less important or more important than you originally thought." Answer categories listed after each benefit were (1) very much more important, (2) somewhat more important, (3) no change in importance, (4) somewhat less important, (5) very much less important, (6) no opinion, (7) do not know, and (8) not applicable.

When IMP respondents report a change in the importance of listed benefits, the change is an *increase* in importance. From 30% to 49% of IMP respondents report an increase in importance for the 11 benefits listed in Table 7.2.

Table 7.2. *Increases in benefits' importance*

Benefit	% Increase
An increase in your library's role as a viable partner in the research enterprise	48.7
Longtime preservation of your institution's digital output	35.0
An increase in the accessibility to knowledge assets such as numeric, video, audio, and multimedia data sets	35.0
Better service to contributors	34.2
Better services to your institution's learning community	34.2
A solution to the problem of preserving your institution's intellectual output	32.5
Exposing your institution's intellectual output to researchers around the world who would not have access to it through traditional channels	32.5
New services to learning communities beyond your institution	32.5
A boost to your institution's prestige	31.7
Capturing the intellectual capital of your institution	30.0
Maintaining control over your institution's intellectual property	30.0

The library's role as a viable research partner makes the biggest jump, registering an almost 50% increase. Overall, 11 of 16 IR benefits register a 30%-or-more increase in importance between planning and implementation, a result that reinforces our idea about the multifaceted nature of IR benefits. Respondents did not use write-in responses to explain their answers to this question, but MIRACLE Project staff will be able to explore respondent answers in subsequent follow-up activities.

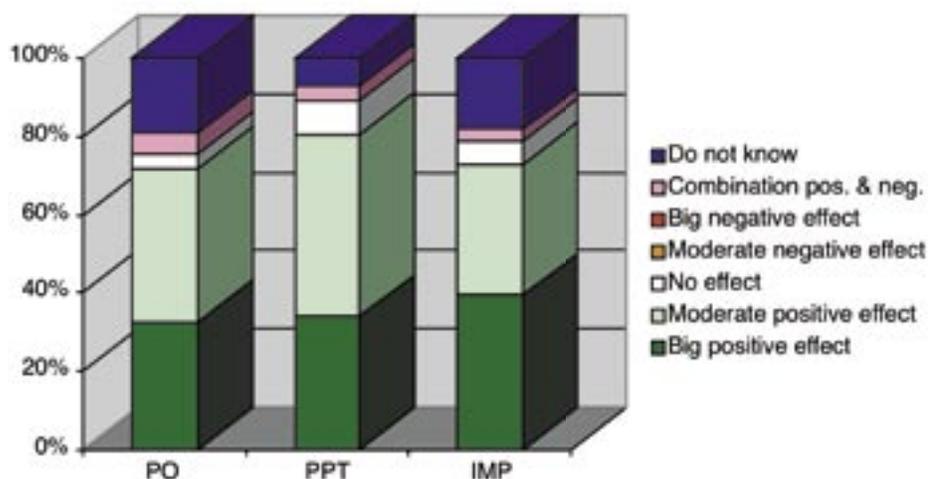
7.2 Building Relationships

Asked to what extent the IR will affect their institution's ability to build relationships with others, such as archives, student services, library systems, and digital asset management systems, PO, PPT, and IMP respondents are overwhelmingly positive (Figure 7.2). No one chooses the answer categories "big negative effect" or "moderate negative effect," and only between 3% and 6% of respondents choose the "combination of positive and negative effects" category. Because of the leading role that libraries take in the IR effort, this question is really a referendum on new relationships forged by libraries as a result of the IR effort.

Larger percentages of IMP respondents (39.4%) check the "big positive effect" than the "moderate positive effect" (33.3%) answer category. Perhaps IMP respondents are especially enthusiastic in their responses to this question because they are starting to see evidence of new relationships as a result of IR implementation.

In the previous question about IR benefits, IMP respondents acknowledge an increased role in the research enterprise (see Table 7.2). Such a role probably comes with the new relationships to which IMP respondents are referring. The nature of these relationships can be explored in this project's follow-up activities, e.g., phone interviews and case studies, when respondents can give open-ended responses to interview questions.

Figure 7.2. Effect on building relationships



7.3 Deploying a Successful IR

Questionnaires asked PO, PPT, and IMP respondents what was likely to inhibit their ability to deploy a successful IR, listed 13 potentially inhibiting factors, and asked respondents to rate those factors. To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very likely; (+1) somewhat likely; (0) no opinion, don't know, or not applicable; (-1) somewhat unlikely; and (-2) very unlikely. They totaled the weights. These results were compiled to rank order all the funding factors. Table 7.3 uses IMP ranks to order factors from top (1) to bottom (13).

Table 7.3. Factors inhibiting the deployment of a successful IR

Top-ranked inhibiting factors (1 to 4)	PO	PPT	IMP
Absence of campus-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints	(5)†	(8)	1
Contributors' lack of knowledge about how they can benefit from IRs	4	1	2
Convincing faculty that the IR will not adversely affect the current publishing model	(10)	(5)	3
Contributors' concerns about intellectual property rights for digital materials	3	4	4
Middle-ranked inhibiting factors (5 to 8)	PO	PPT	IMP
Encouraging faculty to submit digital content to the IR	7	(2)	5
Competing for resources with other priorities, projects, and initiatives	(1)	(3)	6
Making members of your institution's learning community aware of the IR	(12)	(9)	7
Contributors' concerns about the difficulty using the IR system to contribute digital content to the IR	6	7	8
Bottom-ranked inhibiting factors (9 to 13)	PO	PPT	IMP
Supporting all ongoing costs of an operational IR	(2)	(6)	9
Inability of contributors to formulate quality metadata	9	10	10
Difficulties in long-term preservation of digital files	11	11	11
Inadequacy of the IR system's digital preservation capabilities	13	12	12
Lack of on-campus technical expertise in IR systems	(8)	13	13

† Parentheses indicate PO and PPT factors that deviate from IMP top, middle, or bottom ranks.

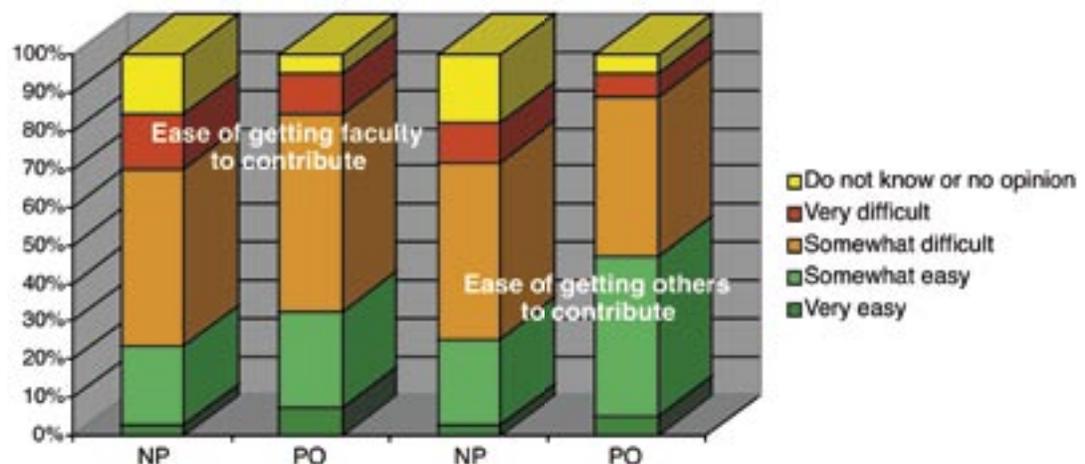
IMP respondents' top-five ranked factors pertain to IR contributors and contributions. In fact, their concern in this regard is pushing them to consider mandating contributions of certain material types. Although PPT respondents are concerned about IR contributors and contributions, they have other priorities, projects, and initiatives that are competing with the IR effort for resources. PO respondents are even more concerned than PPT respondents about sustaining the IR effort, ranking "Competing for resources" and "Supporting ongoing costs of an operational IR" first and second, respectively.

Although there is little agreement among the three respondent types regarding the ranking of inhibiting factors, top-ranked factors for each respondent type reflect their pervasive concerns at their particular stage in the IR effort. IMP respondents have operational IRs. They are concerned about contributors and contributions; in fact, they will even venture to consider mandating deposits of certain materials into the IR. PO respondents do not have a pilot-test or operational IR; thus, their top concern is funding the IR project, then competing with other priorities for resources and securing contributions for the IR. PPT respondents are pilot testing IRs, and thus, they are concerned about securing contributions to the IR.

7.4 IR Contributors

Questionnaires asked NPs and POs, the two respondent types with the least experience with IRs, to speculate on how easy it would be for them to get faculty and other members of their institutions' learning community to contribute to the IR (see Figure 7.3). In view of the concern PPT and IMP respondents have about the success of IRs being connected to IR contributors and contributions (see Table 7.3), we should have invited *all* census respondents to answer this question.

Figure 7.3. Ease of getting people to contribute to the IR



Large percentages of NP and PO respondents think it will be difficult to get faculty to contribute to IRs. NP respondents are more positive than PO respondents about getting other members of their institution's learning community to contribute to IRs; however, about one in eight PO respondents check the "do not know" category.

Questionnaires asked respondents to rate 15 reasons why others would contribute to the IR. For PO and PPT respondents, the question said, "Why do you think members of your institution's learning community will contribute to an IR?" For IMP respondents, the question said, "When planning for an IR, what did you think would be the most important reasons why people would contribute to an IR?" Respondents rated the reasons on a scale from "very important" to "very unimportant."

To simplify results, MIRACLE Project staff assigned weights to response categories as follows: (+2) very important; (+1) somewhat important; (0) no opinion, don't know, or not applicable; (-1) somewhat unimportant; and (-2) very unimportant. The staff totaled the weights. These results were then compiled to rank order all the reasons. Table 7.4 uses IMP ranks to order the top- (1 to 5), middle- (6 to 10), and bottom-ranked (11 to 15) reasons. Parentheses indicate NP, PO, and PPT reasons that deviated from IMP top, middle, or bottom ranks.

Table 7.4. Reasons for contributing to the IR

Top-ranked reasons (1 to 5)	PO	PPT	IMP	Total
To expose the particular scholar's intellectual output to researchers around the world who would not have access to it through traditional channels	3	1	1	1
To boost the particular scholar's prestige	1	2	2	2
To increase the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets	4	3	3	3
To place the burden of preservation on the IR instead of on individual faculty members	5	(6)†*	(7)	4
To solve the problem of preserving your institution's intellectual output	2	(6)T	(10)T	5
Middle-ranked reasons (6 to 10)	PO	PPT	IMP	Total
To expose your institution's intellectual output to researchers around the world who would not have access to it through traditional channels	6	(4)	6	6
To increase citation counts to the particular scholar's oeuvre	8	(5)	(5)	7
To reduce the amount of time between discovery and dissemination of research findings to scholarly communities	7	8	(4)	8
To encourage other scholars to provide open access to their intellectual output	(12)	9	8	9T
To provide maximal access to the results of publicly funded research	9	(11)	10T	9T
Bottom-ranked reasons (11 to 15)	PO	PPT	IMP	Total
To boost your institution's prestige	11	(10)	(9)	11
To increase the library's role as a viable partner in the research enterprise	(10)	12	13	12
To increase citation counts to your institution's intellectual output	13	13	12	13
To contribute to the reform of the entire enterprise of scholarly communication and publishing	14	14	14	14
To reduce user dependence on your library's print collection	15	15	15	15

† Parentheses indicate PO and PPT positions that deviated from IMP top, middle, or bottom ranks.

* T indicates a ranked position that tied another position's weight.

In Table 7.4, respondents give high ratings to reasons that enhance faculty's scholarly reputations and assign responsibility for research-dissemination tasks to others so that faculty can focus on intellectual tasks. Lower-ranked reasons pertain to enhancing the institution's standing.

MIRACLE Project investigators listed the reasons "Boosting the institution's prestige" and "Reforming scholarly communication" because they are prominent in discussions about the ability of IRs to derail the current publishing model (Chan 2004; Crow 2002a; Harnad 2001b). These reasons, however, are not the ones that respondents feel motivate faculty to contribute to IRs; in fact, respondents rank these reasons toward the bottom.

Most census respondents are positive about the importance of all listed reasons. Only the two bottom-ranked reasons have most respondents checking the two "unimportant" response categories.

7.5 Evaluation Methods

Table 7.5 tells the number and percentage of IMP respondents who have used listed methods to assess their IR's success.

Table 7.5. Methods of assessing the IR's success

Methods	IMP	
	No.	%
Tracking number of contributions	27	56.3
Tracking number of users	21	43.8
Tracking number of unique contributors	19	39.6
Tracking number of searches	19	39.6
Conducting interviews with IR contributors	10	20.8
Tracking number of queries	9	18.8
Tracking number of unique IR users	7	14.6
Conducting interviews with IR users	5	10.4
Surveying IR contributors	4	8.3
Surveying IR users	2	4.2

Over half of IMP respondents are tracking the number of contributions to their IRs. Other popular methods are tracking the number of users, unique contributors, and searches. Popular approaches enlist simple counts that the IR system probably produces automatically in periodic management reports. Less popular are interviews with IR contributors (20.8%) or IR users (10.4%) and surveys of IR contributors (8.3%) or users (4.2%). These methods are more intensive, requiring staff to draft data-collection instruments, submit them to institutional review boards for human subjects approval, recruit respondents, collect and analyze data, and communicate results.

Questionnaires should have featured a response category for "Nothing to date" because three IMP respondents wrote us saying they have not collected any data for evaluation purposes.

7.6 Chapter 7 Summary

Census respondents give high ratings to more than a dozen benefits of IRs. In fact, their ratings are so high it is hard to single out one or two benefits as more important than the others (see Subchapter 7.1). Two explanations are given for this finding: (1) IRs have many benefits; or (2) it may be premature to single out any particular benefits because IRs have not yet come into their own.

Questionnaires asked IMP respondents to examine IR benefits a second time, reassessing whether benefits are more or less important now that they are implementing or have implemented an IR. When respondents note a change, the change is an *increase* in importance (see Table 7.2). The library's role as a viable research partner makes the biggest jump; almost 50% of IMP respondents rated this benefit as increasing in importance.

Asked to what extent the IR will affect their institution's ability to build relationships with others such as archives, student services, and library systems, PO, PPT, and IMP respondents are overwhelmingly positive (see Figure 7.2). Hardly anyone chooses "Negative effect" or "Combination of positive and negative effects" categories.

Clarifying what it means to be a viable research partner and exploring the new relationships that the library has established as a result of the IR can be explored in this project's follow-up activities, when respondents can give open-ended responses to interview questions.

Findings about factors that are likely to inhibit their ability to deploy a successful IR reflect the pervasive concerns of PO, PPT, and IMP respondents at their particular stage in the IR effort (see Table 7.3). Because IMP respondents have operational IRs, they are concerned about contributors and contributions. In fact, they will even consider mandating deposits of certain materials into the IR. PO respondents do not have a pilot-test or operational IR; their concerns relate primarily to competing with other projects, priorities, and initiatives for resources at a time when they will soon be acquiring hardware, IR-system software, and the requisite technical expertise. PPT respondents are pilot testing IRs; thus, they are concerned about securing contributions to the IR.

The four top-ranked reasons why census respondents think people will contribute to IRs are connected with enhancing scholarly reputations and offloading research-dissemination tasks onto others (see Table 7.4). Reasons pertaining to reforming the current publishing model figure toward the bottom of the ranked list.

Methods that institutions are using to evaluate IRs usually enlist simple counts that IR systems produce automatically in periodic management reports (see Table 7.5). Less popular are interviews and surveys that require staff to dedicate considerable time and effort to planning, data collection, analysis, and reporting.

8 INSTITUTIONS THAT HAVE NO INVOLVEMENT WITH IRS

Chapter 8 features findings pertaining to institutions that have done no planning for an institutional repository (IR).

8.1 Reasons for No Planning

Participating in the MIRACLE Project's nationwide census are 236 respondents where no IR planning (NP) has been done to date. Dominating the NP respondent type are institutions from the Carnegie Classification of Institutions of Higher Education (CCHE) master's (43.6%) and baccalaureate (33.5%) classes (see Table 2.3).

Questionnaires asked NP respondents to rate 15 reasons why they have not yet done such planning. Table 8.1 tells the percentages of respondents who gave each reason a "very" or "somewhat" important rating.

Table 8.1. Reasons for no planning

Rank	Top-ranked reasons (1 to 5)	% Important
1	Other priorities, issues, activities, etc., are more pressing than an IR	87.2
2	We have no resources to support planning	71.1
3	We want to assess IRs at institutions like our own before taking the plunge	65.5
4	We have no in-house expertise for planning	58.8
5	We want to assess IRs at other institutions generally before taking the plunge	56.1
Rank	Middle-ranked reasons (6 to 10)	%
6	We are waiting for funding to support IR planning	48.2
7	We have no support from our institution's administration	36.3
8	We are waiting to join a consortium, partnership, or group	36.0
9	We doubt members of our institution's learning community will contribute to an IR	33.3
10	We are not convinced that an IR would benefit our institution's learning community	32.9
Rank	Bottom-ranked reasons (11 to 15)	%
11	We have no support from our institution's information technology group	23.4
12	We do not understand or believe in the value or effectiveness of an IR	19.4
13	We will outsource IR services to another institution, consortium, partnership, or group	16.8
14	We do not need an IR	15.4
15	We have no support from our library's administration	9.5

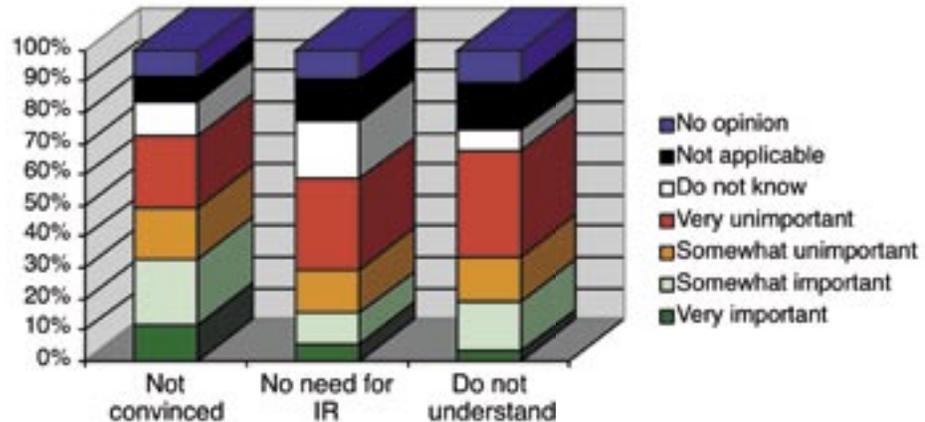
None of the top-ranked reasons rules out these institutions from getting involved with IRs at a later date. Right now, NP institutions appear to have other things on their plates, they have neither the resources nor the expertise needed for IR planning, or they want to assess what comparable institutions have done before taking the plunge.

Reasons ranked 10, 12, and 14 are ones that might preclude academic institutions from becoming involved with IRs. They are ranked almost at the bottom of Table 8.1. Had these reasons been ranked at the top, we might be persuaded that NP respondents have little interest in IRs or do not consider them appropriate for their

institution's learning community. Responses indicate that this is *not* the case.

Figure 8.1 graphs NP respondents' exact answers to these three reasons. Between 40% and 49% of NP respondents consider them "unimportant," and between 33% and 41% are undecided, choosing the answers "Do not know," "Not applicable," or "No opinion."

Figure 8.1. Low-ranked reasons for no IR planning



The no-planning reason getting the least support is "We have no support from our library administration," implying that NP respondents generally do have support from the library administration. Because libraries play such major roles in IR implementation, we do not think that the alternative explanation—NP respondents do not think library support is important for IR implementation—applies here.

Write-in answers give details that are impossible for respondents to express in closed-ended questions. Here is how one NP respondent explained why his or her institution could not take on an IR:

- "Our institution is almost 200 years old and has rich historical resources and an administration that values the resources and is willing to support them financially. We have not yet begun conversations about IR. We have, however, hired our first professionals: museum director, museum registrar, university archivist, and special collections librarian. They are tackling the initial processing of huge collections. We are also completing construction of a \$9 million museum and significantly expanding and renovating the space in the library for archives and special collections."

A handful of respondents write that they have no need for an IR because of their institution's emphasis on teaching over research:

- "Our faculty do not do research as we are a computer-aided design school."
- "We are a very small institution (500 full-time equivalents [FTEs]) with an emphasis on teaching and produce relatively little material of this nature."

Write-in answers reveal two answer categories that MIRACLE

Project investigators should have included on the questionnaires—one about consortia and a second about the issue of IRs never having been raised. Respondent comments about these are as follows:

- “We are in the talking stage internally and are in conversation with a consultant to meet with us and sort out the issues as well as a strategy for eliciting interest in the university. Moreover, our library and IT consortia are interested in a joint endeavor to analyze regional resources and gaps in resources for pursuing institutional repositories and/or consortial repository.”
- “Participating in state-wide library planning effort.”
- “To my knowledge, the issue [of IRs] has never been raised.”
- “There does not appear to be any college discussion or support [for an IR initiative].”
- “This [IR] question has never been discussed on our campus as far as I know.”

Two NP respondents note that they are in the dark about IRs:

- “We do not understand what an institutional repository is.”
- “We are not aware of this whole topic as you obviously are aware by the ‘don’t know’ responses.”

Had large numbers of NP respondents expressed comparable sentiments, MIRACLE Project investigators would have been convinced that NP respondents were disinterested in IR implementation or did not think them appropriate for their institutions. Having encountered few such comments, we conclude that NP respondents simply do not have IR implementation on their agendas right now.

8.2 IR-related Activities

The questionnaires listed 14 IR-related activities, events, or issues that might put NP respondents on the road to an IR and asked them to rate the importance of each item on a scale of from “very important” to “very unimportant.” Table 8.2 tells the percentages of respondents who rated the IR-related activity “very” or “somewhat” important on the road to an IR.

Table 8.2. IR-related activities on the road to an IR

Rank	Top-ranked activities (1 to 5)	% Important
1	How much it costs to implement an IR	90.2
2	How much it costs to maintain an IR	89.6
3	How much it costs to plan for an IR	83.0
4	What institutions comparable to my own are doing with regard to IRs	81.3
5	Whether members of my institution's learning community will use our IR	74.9
Rank	Middle-ranked activities (6 to 10)	%
6	Whether members of my institution's learning community will contribute to our IR	73.9
7	How to interest my institution's administration in IR planning	72.3
8	What is the impetus for IR planning and implementation at institutions comparable to my own	71.6
9	What other institutions generally are doing with regard to IRs	71.4
10	An IR as an accepted "best practice" in the profession	68.2
Rank	Bottom-ranked activities (11 to 14)	%
11	How much it costs to migrate to a new IR	63.8
12	What is the impetus for IR planning and implementation at other institutions generally	59.8
13	How to interest a consortium, partnership, group, library network, etc., in IR planning	48.3
14	How to interest an institution(s) in partnering with us on an IR	37.1

First, NP respondents are concerned about the *costs* of IRs. Next, they want to know what institutions *comparable* to their own are doing with regard to IRs. They then want to know whether members of their institution's learning community will *contribute to* and *use* their IR. Finally, they want to know how to interest their institution's administration in IR planning.

Less than half of NP respondents give "important" ratings to only the bottom two Table 8.2 activities. Both these activities address partnering with other institutions for IR services. Clearly, NP respondents in the MIRACLE Project census prefer to go it alone in terms of IR services.

Because NP respondents give high ratings to almost all IR-related activities, they are surprisingly very favorably inclined to IRs. This may be because of how we invited people to participate in the census. We performed the electronic version of the salesperson's "cold call"; that is, we sent prospective respondents e-mail messages with a substantive phrase in the "SUBJECT" line announcing our IR census and asked them to participate. Most likely, the people who responded to our e-mail message are interested in IRs and are thus more likely than others to open and read such a message and to respond positively about IRs on their questionnaire.

Questionnaires asked NP respondents to choose one or more events that would have to happen for IR planning to begin at their institutions. Table 8.3 tells the percentages of respondents who chose the listed event.

Table 8.3. What would have to happen for IR planning to begin at your institution?

Rank	Top-ranked events (1 to 4)	% Important
1	We receive funding from our institution's administration	66.1
2	Successful IR demonstration projects at a comparable institution	54.7
3	We receive approval from our institution's administration	52.1
4	We are convinced that our institution's learning community would contribute to it	43.2
Rank	Middle-ranked events (5 to 7)	%
5	We receive additional personnel resources to support planning	36.4
6	We receive approval from our institution's information technology group	34.3
7	We reassess our institution's current priorities, issues, and activities	33.9
Rank	Bottom-ranked events (8 to 12)	%
8	Successful IR demonstration projects at other institutions generally	19.5
9T	Contracting for IR services from another institution, consortium, or group	18.2
9T	We receive approval from our library's administration	18.2
11	We receive funding from our institution's information technology group	17.4
12	We receive funding from our library's administration	12.7

To initiate IR planning, NP respondents need approval from their institution's administration and funding that includes support for the personnel to undertake the project. Respondents also want evidence of successful IR projects at comparable institutions. They are *not* interested in evidence of such projects at institutions unlike their own and are *not* interested in partnering for IR services. Because most NP respondents come from master's and baccalaureate institutions, they want to be convinced that IRs at institutions awarding these degrees are successful in terms of technical implementation, securing contributions to the IR from the local learning community, and system use by the local community and beyond.

NP respondents' open-ended responses to this question reveal five themes: (1) the pressing nature of other priorities, issues, activities, etc.; (2) the need for resources to begin planning; (3) waiting for a consortium; (4) low levels of research at the institution; and (5) raising the issue of IRs at their institution. Here are examples of each:

1. Other priorities, issues, and activities:

- "[The] IR is one of the university librarian's hot topics for his 10-year plan for content services. For now, however, other issues are more pressing."
- "Achieve stability in the administration of the institution."
- "Many of these requirements are already in place; delay is just because we have other projects higher on the priority list."

2. Resources needed:

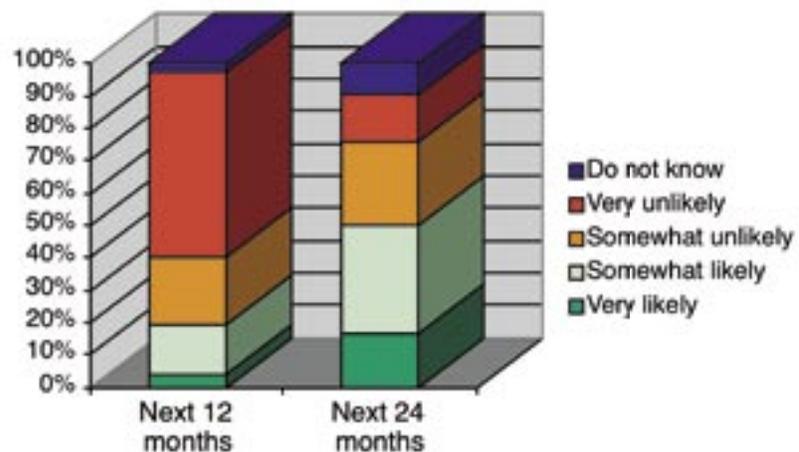
- "We have appropriate, trained, skilled personnel who are committed to overseeing the project."
- "We [must] receive staff support [for the IR project] from [the] information technology group."
- "We find some outside funding."

- “We have limited staff to engage in the planning, promotion, education, etc. Funding may eventually be a problem, but until planning is done, I can’t say for certain.”
3. Waiting for a consortium:
 - “Move by the state board of regents to develop a system-wide IR.”
 - “Consortium, could move toward a statewide IR site.”
 4. Low research levels:
 - “This is not a research institution, hence IR planning is not a high priority here.”
 - “We need to publish more.”
 5. Raising the issue:
 - “Someone at the administrative level would need to embrace an IR as a goal for our institution—right now, an IR seems to be completely off radar.”

8.3 Likelihood of IR Planning

Two questions asked NP respondents about the likelihood of future IR planning. The first question asked them about such a likelihood in the near future (next 12 months), and the second asked them about this in the medium term (next 24 months). Figure 8.2 shows the results.

Figure 8.2. Likelihood of IR planning



Less than 20% of NP respondents are likely to start IR planning in the next 12 months. Just under 50% are likely to start such planning within 24 months. An operational IR is a distinct possibility for many of the NP institutions participating in the MIRACLE Project census.

8.4 Chapter 8 Summary

The top-ranked reason why NP institutions have done no IR planning to date is the pressing nature of other priorities, issues, and

activities (see Table 8.1). None of the other top-ranked reasons rules out these institutions from eventually getting involved with IRs. Had reasons such as “We do not need an IR” or “We do not understand or believe in the value or effectiveness of an IR” been top ranked, then we would question whether NP institutions would get involved with IRs in the short- to medium-term future.

Right now, NP institutions appear to have other things on their plate, they have no resources or expertise for IR planning, or they want to assess what others are doing before taking the plunge.

Asked to rate the importance of 14 next steps on the road to an IR (see Table 8.2), NP respondents give the highest ratings to three steps pertaining to costs:

- “How much it costs to implement an IR”
- “How much it costs to maintain an IR”
- “How much it costs to plan for an IR”

After learning about costs, NP respondents want to know what institutions comparable to their own are doing with regard to IRs. Then they want to know whether members of their institution’s learning community will contribute to and use the IR. Finally, they want to know how to interest their institution’s administration in IR planning. They are *not* interested in partnering with other institutions. When asked about learning from other institutions, they want to know about the IR-implementation experience of *comparable* institutions, meaning master’s and baccalaureate institutions, because these are the majority of NP institutions responding to the MIRACLE Project census.

Underlying the high ratings NP respondents give to all but a handful of the next steps on the road to IR planning is a favorable inclination toward IRs. This may be because of how MIRACLE Project staff invited people to participate in the census, performing an e-mail version of cold calling. Most likely, the people who responded to our e-mail message are interested in IRs and are thus more likely to open and read a message about them and, eventually, respond positively about IRs on their questionnaires.

To initiate IR planning (see Table 8.3), NP institutions need approval from their institution’s administration and funding that includes support for personnel. They also want evidence of successful IR projects at comparable institutions; again, this means master’s and baccalaureate institutions. They are *not* interested in evidence of such projects at institutions unlike their own and are *not* interested in contracting for IR services. Several NP respondents wrote open-ended responses that reveal these themes: (1) the pressing nature of other priorities, issues, activities, etc.; (2) the need for resources to begin planning; (3) waiting for a consortium; (4) low levels of research at the institution; and (5) raising the issue of IRs at their institution.

Less than 20% of NP respondents are likely to start IR planning in the next 12 months. The percentage increases to just under 50% for a beginning start date in 24 months.

Generally, NP respondents in the MIRACLE Project census are

favorably inclined toward IRs. They come from master's and baccalaureate institutions where research may not be as important as teaching and service. These respondents recognize the importance of IRs for their institutions and for educational institutions in general. Because they have other priorities right now, they are content to take a wait-and-see approach, that is, monitoring whether IRs at institutions like their own have been successful in technical implementation, whether members of their learning communities will contribute to IR, and whether they will use the IR. Planning for the cost of IR implementation, finding staff with the requisite expertise, and broaching the issue of IR implementation with their administration are important issues for NP respondents.

9 DISCUSSION OF CENSUS FINDINGS

Chapter 9 discusses census findings, specifically, the sleeping beast of demand for institutional repositories (IRs) from master's and baccalaureate institutions, findings that confirm those of previous surveys, and findings that build on our knowledge of IRs. It concludes with observations on issues pertaining to IRs that will persist long after the MIRACLE Project terminates.

9.1 The Sleeping Beast of Demand for IRs from Master's and Baccalaureate Institutions

No planning (NP) institutions are the largest respondent type in the MIRACLE Project's nationwide census, accounting for 52% of respondents (see Figure 2.1). Planning only (PO) institutions are the second-largest respondent type in the census, accounting for 21% of respondents (see Figure 2.1). Dominating both NP and PO respondent types are institutions from the Carnegie Classification of Institutions of Higher Education (CCHE) master's (43.6% and 34.8%, respectively) and baccalaureate (33.5% and 31.5%, respectively) classes (see Table 2.3).

Despite their prevalence (56.6%) in the population of institutions of higher education (IHEs) in the United States (see Figure 2.3), master's and baccalaureate CCHE institutions are not where IR activity is happening. To date, the story of IRs in U.S. academic institutions has been written by the research universities CCHE class. Although research universities represent only 7.9% of IHEs in the U.S. (see Figure 2.3), they are the majority (62.5%) of IMP institutions in the MIRACLE census, that is, institutions where IRs have been implemented (see Table 2.3). Previous IR surveys (Bailey et al. 2006; Shearer 2004) have been limited to members of the Association of Research Libraries (ARL) in the United States and of the Canadian Association of Research Libraries (CARL)—types of libraries that are typical of the research universities CCHE class. The Coalition for Networked Information (CNI), another recent surveyor, is sponsored

by ARL, and CNI's survey of 81 liberal arts colleges bearing CNI consortial membership reveals that only 6% have an operational IR (Lynch and Lippincott 2005). Two prominent research universities, Massachusetts Institute of Technology (MIT) and Cornell University, have been involved in the development of the popular DSpace and Fedora IR systems, respectively. Case studies (e.g., Smith et al. 2003; Rogers 2003; Walters 2006; Baudoin and Branschofsky 2003) focus on research library and research university experiences with IRs. IRs are a recent phenomenon and they are happening at *research universities*.

The MIRACLE Project census has uncovered a sleeping beast of demand on the part of master's and baccalaureate universities and colleges regarding IRs. Respondents at these institutions want to know about the IR experiences of master's and baccalaureate institutions generally (see Tables 8.1, 8.2, and 8.3). They also want to learn about their peers' experiences with IR costs, required technical expertise, funding the IR effort, whether the local learning community will contribute to and use the IR, and raising the issue of IRs with their institution's central administration.

MIRACLE Project questionnaires ended with this question: "How can the MIRACLE Project assist you regarding IRs?" Many NP and PO respondents asked us particularly about the small and mid-size college and university experiences with IRs. Samples of such responses include the following:

- "Tell us stories about how small institutions made their IRs a reality. We really need models and realistic next steps." (PO respondent)
- "Best practices, identification of institutions like ours [a small liberal arts college in the middle Atlantic states] who have succeeded, formation or information about collaborative groups who have (or will have) a shared IR that we can join." (NP respondent)
- "I believe that a full-fledged IR is beyond our capabilities at this point, but would be interested in continuing to hear about developments in this area, especially in small universities." (NP)
- "Provide examples of what other small to mid-size public universities are doing with IRs." (NP)
- "Would love to see models in a small, liberal arts college environment, particularly for consortial opportunities." (NP)
- "We are always interested in what our peer institutions [mid-size public midwestern universities] are doing." (NP)
- "Providing information about what comparable institutions [small liberal arts colleges in the Central Plains states] are doing ..." (NP)
- "By publicizing success stories from institutions similar in size and mission to ours [small private midwestern college]." (NP)
- "Provide more information about IRs, their benefits, and the resources needed to establish and maintain them. We are a small, underfunded undergraduate [private Southern] college just struggling to fund basic needs." (NP)

NP respondents participating in the MIRACLE Project are surprisingly positive about IRs. Very few are totally in the dark in terms

of what IRs are and whether they have relevance for their institutions (see Figure 8.1). Slightly less than 50% of NP respondents may start IR planning within the next 24 months (see Figure 8.2).

The positive attitude that this project's NP respondents have about IRs may be the result of how MIRACLE Project staff invited people to participate. We performed the electronic version of cold calling, that is, we sent prospective respondents e-mail messages with a substantive phrase in the "SUBJECT" line announcing our IR census and asked them to participate. Most likely, the people who responded to our message are interested in IRs and are more likely to read and respond to such a message, and, eventually, to respond positively about IRs on their questionnaire.

MIRACLE Project investigators identified more themes in NP respondents' answers to our question "How can the MIRACLE Project assist you regarding IRs?" These themes are (1) learning about IRs generally, (2) learning the details and specifics of IRs, (3) best practices, (4) benefits of IRs, (5) securing funding for IRs, (6) encouragement and advocacy, (7) opportunities for partnerships, and (8) learning about IRs from completing the MIRACLE Project's questionnaire. Table 9.1 lists a few remarks for each of these themes. Many more examples could be enumerated in this table, and most remarks cut across two or more themes.

The high level of interest in IRs is an opportunity for other-than-research-universities to share their stories about IRs with an audience that is craving information. It is also an opportunity for the MIRACLE Project to focus on other-than-research-universities in subsequent project activities, where the need is greatest and where the widest gap in our knowledge about IRs exists.

Table 9.1. How the MIRACLE Project can assist NP respondents regarding IRs

Theme	NP respondent's remark	Institution detail
(1) Learning about IRs generally	"I'm still learning what IRs are and how we might think about starting one ourselves. Any information on those topics would be useful to me at this point."	Special focus professional school in the Southwest
	"Continue disseminating information. Review and publicize tools, especially those for institutions with limited technical support and funding."	Small private liberal arts college in the Southeast
	"I think your study itself will be valuable."	Western master's university
	"Send us survey results. Connect us to institutions like us who are considering IRs."	Small private liberal arts college in the Central Plains states
(2) Learning the details and specifics of IRs	"Marketing materials, potential benefits and liabilities ... The whole administrative impact. From the smallest size institution, this is more than just adding a service; it could relate to a huge percentage of extremely tight resources. Erase the FEAR of costs."	Small private church-affiliated liberal arts college in the Great Lakes area
	"Information and assistance on coping with copyright issues associated with an IR; promotional material and arguments to convince faculty and the learning community to participate in and support an IR."	Mid-size public research university in the Mountain West
(3) Best practices	"Provide 'best practices' for an institution of my size. Offer guidelines for partnering with other institutions."	Small public baccalaureate university in the Mountain West
	"We'll be interested in the procedures developed by others and what current best practices are at the time we're ready to start."	Small private master's university in New England
(4) Benefits of IRs	"We are in the midst of an institution-wide reassessment. The benefits listed in question 6a reflect many of the values we would like to incorporate in our plans for the future [and] ... to the academic community as a whole."	Small private master's university in a large northeastern city
	"Provide concise examples and talking points of benefits and successes from IRs for use to gain campus and administrative support."	Mid-size public master's university in a Central Plains state
(5) Securing funding for IRs	"Figure out a way so that the top administration would want to fully fund such an operation. It really requires a lot of talented labor to input and maintain."	Small private research university in the Northeast
	"Provide more information about IRs, their benefits, and the resources needed to establish and maintain them. We are a small, underfunded, undergraduate college just struggling to fund basic needs."	Small private church-affiliated baccalaureate college in the South
(6) Encouragement and advocacy	"This is really low on the radar right now. Just being there in the future is all I could ask at this time."	Small technical [special-focus] institution in the Northern Plains states
	"If you can help me wake people up to the potential of an IR over the din of all the other challenges of an institution like ours, that would be great."	Small private religious-affiliated master's university in the Northeast
(7) Opportunities for partnership	"Suggest some consortial models we could investigate."	Small private liberal arts college in the Southeast
	"Would love to see models in a small, liberal arts college environment, particularly for consortial opportunities."	Small private master's university in the Southeast
(8) Learning about IRs from the questionnaire	"You have already identified the issue for me. I will leave it simmering on the back burner until I see more interest within the faculty and the library community in general."	Small, private, church-affiliated liberal arts college in the Central Plains states
	"Having this kind of in-depth survey to use as background information and ammunition will help spark the whole planning and implementation."	Mid-size midwestern doctoral university in a Central Plains state
	"Presenting the range of questions we should be thinking about, so just taking this survey has been educational."	Major military academy

9.2 Verifying Previous Survey Findings Pertaining to Institutions Involved with IRs

Findings from the MIRACLE Project census verify previous survey findings pertaining to institutions with operational IRs (see Appendix F). Table 9.2 summarizes the most important of these findings. The Executive Summary is also comprehensive in its enumeration of MIRACLE Project census findings.

Table 9.2. Previous survey findings verified in the MIRACLE Project census

Finding	Report references
Research universities lead in the implementation of IRs.	Table 2.3 and Subchapter 2.2
Master's and baccalaureate institutions lag far behind in the implementation of IRs.	Table 2.3 and Subchapter 2.2
Libraries play a leading role in planning, pilot testing, and implementing IRs.	Table 2.4 and Subchapter 2.3; Table 2.5, Figure 2.5, and Subchapter 2.5; Figure 2.6, and Subchapter 2.6; Table 3.1
Committee membership becomes increasingly less inclusive as the IR project progresses from pilot testing to implementation.	Figure 2.5 and Subchapter 2.5
The number of staff involved in the IR effort decreases from the planning and pilot-testing stage to the IR implementation stage.	Figure 2.4 and Subchapter 2.4
Libraries bear the brunt of the cost of the IR.	Table 3.1 and Subchapter 3.1
A typical approach to funding the IR is absorbing its cost in routine library operating costs.	Table 3.1 and Subchapter 3.1
Staff and benefits costs dominate the budget for IR.	Figure 3.1 and Subchapter 3.2
Pilot testing IR-system software is an important investigative activity.	Table 4.3 and Subchapter 4.3
Institutions' preferred IR-system software for both pilot testing and implementation is DSpace.	Table 5.2 and Subchapter 5.2
Most IR staff modify their IR-system software.	Figure 5.1 and Subchapter 5.3
Both pilot-test and operational IRs are very small.	Figure 6.1, Table 6.1, and Subchapter 6.1
Dominating pilot-test and operational IRs are the traditional products and by-products of the research enterprise.	Table 6.1 and Subchapter 6.1
Operational IRs contain a wide range of text, numeric, and multimedia files.	Table 6.1 and Subchapter 6.2
Except for PDFs, institutions with operational IRs do not guarantee file formats in perpetuity.	Table 6.2 and Subchapter 6.4
The major contributors to operational IRs are faculty or graduate students.	Table 6.4 and Subchapter 6.5.2
Recruiting digital content for the IR is difficult.	Figures 6.7 and Subchapter 6.5.4; Table 7.3 and Subchapter 7.3; Figure 7.3 and Subchapter 7.4
IR staff working one-on-one with early adopters is a successful method for recruiting IR content.	Figure 6.5 and Subchapter 6.5.3
IR staff may consider institutional mandates that require members of their institution's learning community to deposit certain document types in the IR.	Subchapter 6.5.4; Table 7.3 and Subchapter 7.3
For IR staff, top-ranked benefits of IRs are institution based.	Table 7.1 and Subchapter 7.1
Evaluation methods to date are limited to simple counts that most IR systems produce automatically in management reports.	Table 7.5 and Subchapter 7.5

9.3 Building on Our Knowledge of IRs

MIRACLE Project findings that build on our knowledge of IRs are featured in this subchapter. Because the questionnaire method does not allow MIRACLE Project investigators to gain insight above and beyond closed-ended responses, we must sometimes speculate on the reasons for these findings. Subsequent project activities (e.g., phone interviews, case studies, follow-up electronic mail correspondence, and quasi-controlled experiments with IR users) will determine whether our speculation, arguments, and reasoning are on target.

1. Except for the library director, the key people who would have to be very active to initiate an IR effort where none is under way are external to the library. (See Table 2.4 and Subchapter 2.3, and Table 8.3. See also Table 4.1 and Subchapter 4.1.)

Librarians are especially active in the IR effort at PO, PPT, and IMP institutions, serving on planning and advisory committees, pilot testing software, recruiting content, identifying early adopters, etc. At NP institutions where no IR effort is under way, the library director takes the lead, inquiring about funding from the provost and technical expertise from the chief information officer (CIO) and learning about the faculty's interest in making contributions and urging their students to make contributions.

Related to this finding are the important investigative activities that IR staff undertake. Especially important to PO respondents is demonstrating operational IRs to their institutions' decision makers (see Table 4.1). Because PO respondents are in the early stages of the IR initiative, they want those who will ultimately be making the decision about their institution's IR effort and, possibly, giving financial support for the IR project, to understand the basic concept of IRs. Demonstrating IRs makes them more tangible to decision makers so they may be more favorably inclined to the IR initiative in terms of both funding and rhetoric.

2. Archivists are less prominent in the IR effort than expected. (See Tables 2.4 and 2.5, Figures 2.5 and 2.6, Table 3.1 and Subchapter 3.1, and Subchapters 2.3, 2.4, 2.5, and 2.6.)

Despite the inclusion of "archivists" response categories on questionnaires, archivists figure in the middle when querying respondents about the positions of people involved in the IR effort (see Table 2.4), leading the IR effort (see Table 2.5), and serving on IR committees (see Figure 2.5). Between the planning through implementation phases of the IR effort, archivists' responsibility for the IR appears to diminish (see Figure 2.6). Archivists are also not expected to bear the burden of funding the IR project (see Table 3.1).

MIRACLE Project investigators have no census data that would help explain the marginalization of the archivist with respect to IRs. There may be merit to Crow's (2002a) observation that the IR competes with the university archives (see Appendix F4).

IRs could benefit from archivists who are experts in collection

building and disposition. The type of one-on-one collection development and content recruitment now being carried out by librarians to populate IRs is exactly the type of field work that archivists have done for decades. Closely related to this type of content recruitment is archival appraisal, which is a different type of collection analysis for librarians and pushes their skill set into the archival arena.

In future studies, the relationship between IRs and archivists deserves further investigation to shed light on the reasons for archivists' limited participation in IR efforts.

3. Staff involved with the IR effort have voracious appetites for information about IRs, especially information pertaining to successful implementations at institutions like their own. (See Table 4.1 and Subchapter 4.1, Tables 8.1 and 8.2, and Subchapters 8.1 and 8.2.)

The tables cited above only touch the surface in terms of what respondents want. The wide range of their interests and needs is demonstrated in their answers to the final question on MIRACLE Project questionnaires, "How can the MIRACLE Project assist you regarding IRs?" Because enumerating their specific requests in their own words would be too lengthy for this report, we use Table 9.3 to characterize them.

Table 9.3. Characterizing respondents' requests for more information about IRs

Respondents want information on	NP	PO	PPT	IMP
MIRACLE Project census findings	X	X	X	X
Successful IR implementations especially at institutions like their own	X	X	X	X
Best practices	X	X	X	X
State-of-the-art regarding IRs	X	X	X	X
Examples*	X	X	X	X
MIRACLE Project census data		X	X	X
Costs and budgets	X	X	X	
Case studies especially at institutions like their own		X	X	
Written policies		X		X
MIRACLE Project as a clearinghouse for all information about IRs	X	X		X
Joining a consortium or partnership	X	X		
Benefits of IRs	X	X		
Compelling arguments for an IR in institutions like their own	X	X		
Software reviews	X	X		
Grant funding opportunities		X		
MIRACLE Project questionnaires		X		

* Examples cited by one or more respondent types: written policies, procedures, consortial agreements, permissions templates, intellectual property agreements, requests for proposals, benchmarks, models, collection development policies, IR-system software checklists (especially for open-source alternatives), comparative analyses of IR-system software products.

Respondents from PO institutions are especially demanding and the most articulate about their demands. In fact, the majority of examples come from PO respondents.

4. The needs assessment is not as important as other IR investigative activities. (See Table 4.1 and Subchapter 4.1, Figure 4.1, and Subchapter 4.2.)

The needs assessment ranks in the middle of a dozen investigative activities (see Table 4.1). The majority of MIRACLE Project respondents do not conduct such an assessment (see Figure 4.1); in fact, between 5% and 12% of PO, PPT, and IMP respondents do not even know whether their institution has conducted a needs assessment. Learning about successful IR implementations at comparable institutions ranks head and shoulders above the needs assessment. Table 9.3 lists other information that IR planners and implementers would like in hand.

5. The next steps for IR planners and pilot testers will be to continue, not terminate, their institution's IR effort. (See Table 4.4, Figure 4.5, and Subchapter 4.4.)

The next step for PO respondents is to widen the scope of their IR investigations, and, for PPT respondents, to implement an IR-system software package (see Table 4.4). Terminating the IR project is another logical next step for respondents in these two groups, but only about 10% of PO and PPT respondents in the MIRACLE Project census plan to do so (see Figure 4.5).

6. Waiting for a consortium is a viable alternative for a small minority of institutions interested in IR services. (See Table 4.1 and Subchapter 4.1, Table 4.4, Figure 4.5 and Subchapter 4.4, Subchapter 5.4, Table 8.1 and Subchapter 8.1, and Subchapter 2.6.)

When MIRACLE Project investigators included references to consortia, partnerships, networks, or groups in response categories, few respondents chose these categories. When we failed to include such references, a handful of respondents volunteered them in write-in responses. Asked directly about joining a consortium, respondents rank it in the middle of a pack of reasons why they have not yet begun IR planning (see Table 8.1). Respondents involved with IR planning only or not involved with IRs express their interests in consortia in the final question on MIRACLE Project instruments about how the Project could help them. For example:

- "We are in the process of investigating IR systems and are in talks with other colleges about our digital needs. A consortial agreement for an IR system would be ideal." (PO respondent at a small private liberal arts college in a Great Lakes state)
- "Provide information about collaboratives, either within a consortium, a system, or amongst institutions with similar needs." (PO respondent at a mid-size master's university in a northern Great Lakes state)
- "Offer guidelines for partnering with other institutions." (NP re-

spondent at a small public baccalaureate university in the Mountain West)

- “Best practice, identification of institutions like ours who have succeeded, formation or information about collaborative groups who have (or will have) a shared IR that we can join. We see a shared system as one of the more viable options.” (NP respondent at a small private liberal arts college in the central Atlantic states)
- “Would love to see models in a small, liberal arts college environment, particularly for consortial opportunities.” (NP respondent at a small master’s university in the Southeast)

7. About one-quarter of institutions pilot testing or implementing an IR have two or more IRs available to their institution’s learning community. (See Table 5.1 and Subchapter 5.1.)

Perhaps MIRACLE Project respondents are counting the academic departments and research units that have launched IR-like software to preserve, exchange, and distribute research and teaching objects among themselves, colleagues at other schools, and Web searchers generally. These could also be subject-based repositories that prefer to maintain their subject focus rather than IRs that reflect the encyclopedic nature of liberal arts colleges and research universities. What will happen at institutions with multiple IRs? Will they join forces and consolidate their efforts? What are the forces for and against centralization? Is it advantageous for multiple IRs at a single institution to prosper?

8. The availability of additional commercial options for IR-system software may enable more institutions to get involved with IRs especially at the many master’s and baccalaureate institutions where IR implementation is uncommon. (See Table 5.2 and Subchapter 5.2, Table 4.3 and Subchapter 4.3, and Table 8.1.)

Although the open-source DSpace and Fedora systems are the most popular IR systems for pilot testing and implementation (see Table 5.2), they require systems staff to program, profile, and deploy. An important benefit of pilot testing is to develop the requisite technical expertise for system deployment (see Table 4.3). NP institutions in the MIRACLE Project census have neither the resources nor in-house expertise to support IR planning (see Table 8.1). These institutions could benefit from commercial vendors who install the system and train on-site staff in system management and maintenance. Institutions involved with IR planning only or not at all involved with IRs express their interest in commercial vendors in the final question on MIRACLE Project instruments about how the project could help them. For example:

- “Models of IRs that are managed at the institution and those that are managed by vendors, for example, Digital Commons.” (PO respondent at a small public master’s university in the Southeast)
- “Review and publicize tools, especially those for institutions with limited technical support and funding.” (NP respondent at a small baccalaureate college in the Southeast)

- “Currently investigating platforms and working with a consortium to get RFPs ... We are still planning. Would help to know more about commercial alternatives.” (PO respondent at a small baccalaureate college in a Central Plains state)
 - “Conduct some measure of comparative analyses of IR product offerings or software that would allow a single, inexperienced institution to expedite or focus its own analysis.” (PO respondent at a small master’s university in the Pacific Northwest)
9. IR-system functionality is satisfactory but the user interface including controlled vocabulary searching and authority control needs serious reworking. (See Table 5.3 and Subchapter 5.3.)

Census respondents give high ratings to IR-system functionality for browsing, searching, and retrieving digital content; mediocre ratings to the user interface; and dead-last ratings to controlled vocabulary searching and authority control (see Table 5.3). A user interface that impedes retrieval will send most users packing, that is, switching from the IR to one of the hundreds of different databases that research institutions where most IRs are deployed offer their learning communities. Another option is to switch to Google with its popularity-based retrieval that does a good job ranking relevant retrievals at the top. People who search online systems conform to the *principle of least effort*, “The design of any ... information system should be the system’s ease of use ... If an organization desires to have a high quality of information used, it must make ease of use of primary importance” (Rosenberg 1966, 19).

At this early point in the development and deployment of IRs, few people have searched these systems. Now is the time to make user-interface improvements before too many users have negative experiences and abandon them altogether.

10. Improve preservation functionality in IRs. (See Appendix F1, Table 5.4 and Subchapter 5.4, and Table 6.2 and Subchapter 6.4.) Long-term preservation of digital materials figures prominently in Clifford Lynch’s (2003) definition of IRs (see Appendix F1). If universities, through their IRs, are going to replace the current publishing paradigm, the ability to maintain the documents (in whatever format or medium) over time is required. The promise of the IR, then, is not only to maintain the viability of the byte stream of these materials but also to support technologies that make a variety of file formats accessible over the long term. If one agrees with Lynch’s definition, every IR must become a trusted digital repository (RLG 2002).

Except for PDF files (see Table 6.2), today’s IR systems make few promises about guaranteeing digital file formats in perpetuity. The top reason for census respondents migrating to new IR-system software is greater capacity for handling preservation (see Table 5.4). IR systems must improve their preservation functionality. At the least, such an improvement fulfills a key reason for the very existence of IRs.

11. Institutions do not need policies written in stone at the public launch of their IR. (See Figure 6.2 and Subchapter 6.3.)

At least 60% of census respondents with operational IRs report implemented policies for (1) acceptable file formats, (2) determining who is authorized to make contributions to the IR, (3) defining collections, (4) restricting access to IR content, (5) identifying metadata formats and authorized metadata creators, and (6) determining what is acceptable content (see Figure 6.2). There are many more policies for which these institutions report *drafted* policies or *no* policies at all.

It may be not necessary for all IR policies to be in place at the time of the public launch of an institution's IR. Taking a wait-and-see attitude, evaluating what transpires after a period of time, then firming up existing policies and implementing new ones as needed may be the most expedient course. Here is advice from a respondent whose institution has an operational IR:

- "Halfway through [completing this questionnaire], I realized that it wasn't going to help me at all and that it would only serve to let the timid think that they had to have all of their eggs in the basket before they tried anything. JUST DO IT!" (IMP respondent from a mid-size master's public university in the Midwest)

12. The IR helps libraries build new relationships. (See Figure 7.2 and Subchapter 7.2, and Table 7.2.)

Asked to what extent the IR will affect their institution's ability to build relationships with others such as archives, student services, digital asset management systems, etc., PO, PPT, and IMP respondents respond overwhelmingly positively (see Figure 7.2). Because of the leading role that libraries take in the IR effort, this question really pertains to the new relationships that are a result of the library's involvement with IRs. MIRACLE Project staff can explore the nature of these relationships in this project's follow-up activities, e.g., phone interviews and case studies, when respondents can give open-ended responses to interview questions and interviewers can probe deeper into fruitful areas. A preview of potential findings in this regard comes from IMP respondents who credit the IR with increasing the library's role as a viable partner in the research enterprise (see Table 7.2).

13. To what extent is the impetus for the IR coming from faculty, staff, and students? (See Subchapter 4.2, and Subchapters 7.1 and 7.2.)

The ARL SPEC Kit survey reports "38% of implementers and 47% of planners were responding to requests for an IR from faculty, staff, and students" (Bailey et al. 2006, 25). The MIRACLE Project census did not question respondents directly about the impetus for an IR; however, it did question them extensively about benefits of IRs (see Subchapters 7.1 and 7.2) and respondents failed to volunteer write-in responses that mentioned faculty, staff, and students in this regard.

A few comments that IMP respondents volunteered in response to a question about the importance of conducting a needs assessment

describe faculty interest in IRs that helped start the IR project:

- “Our assessment was more dynamic and ongoing ... it involved response to innovative faculty requests and ongoing outreach from librarians regarding changes in scholarly communication practices ... ”
- “Our former Dean of Faculty was particularly interested in DSpace and secured funding ... to support its use here.”
- There was no needs assessment but the IR was very much faculty driven. Leadership was taken by the University Library Council (a senate-provostial advisory group) that pushed the agenda and prepared the report that led to provost funding and support.”

The difficulty that IR staff are experiencing getting contributions from members of their institution’s learning community may be evidence contradicting the faculty, staff, and student impetus for IRs. In subsequent project activities, MIRACLE Project staff will look for evidence of such an impetus. They will also look for evidence of “peer pressure” as the impetus for an IR.

9.4 Observations on Long-term Issues Pertaining to IRs

MIRACLE Project census findings and the project’s follow-up activities will be able to explore the seven points discussed in this subchapter, but they will not be able to answer them conclusively. Definitive answers are possible only for those who hold a mirror on the future. The passing of time, convergence of events, advances in technology, and the inevitable march of human progress will eventually reveal the future.

1. Is it too early in the evolution of IRs to single out one particular benefit? (See Table 7.1, Figure 7.1, and Subchapter 7.1, Table 7.4 and Subchapter 7.4, and Table 7.5.)

Asked to rate a list of 16 benefits of IRs, census respondents give very high ratings to all but two (see Figure 7.1 and Table 7.1). Instead of having a couple of benefits that stand head and shoulders above the others, IRs may have many benefits. Or it may be premature for one or two benefits to rise above the others because IRs have not yet come into their own. It may be prudent to give IRs a half decade or so to become commonplace in all types of educational institutions and to then pose this question again to the same audience. One benefit may rise above the others.

Because the MIRACLE Project census was limited to IR staff, we could ask them only why they thought people would contribute to IRs (see Table 7.4). They gave high ratings to reasons connected with enhancing faculty scholarly reputations and passing on research-dissemination tasks to others so that faculty can stay focused on intellectual tasks. Lower-rated reasons pertained to enhancing the institution’s standing.

Future surveys should ask IR contributors directly about why they contribute to IRs and the benefits they receive from their contri-

butions and compare results from IR staff regarding benefits.

Today's IRs do not yet have a dedicated corps of end users. No studies have been conducted to determine why people search IRs and whether they find relevant materials of interest. Very few institutions at which IRs are operational are collecting much more than counts of users, unique contributors, and searches (see Table 7.5). Future user studies should question users about IR benefits. IR staff who have done no IR planning or are just beginning to plan will benefit from anecdotes about how IR content helps researchers, for example, enabling a scholar to examine unique primary materials or putting an undergraduate working on a senior thesis project in touch with a scientist who supplied her with data files for analysis. Institution staff will especially use these stories to convince their decision makers to support an IR effort. Here is what they have to say in this regard:

- "If our president and provost were more aware of the benefit of an IR to the scholarly community and for access to the institution's historical record, I would be more likely to meet with success. At this small institution, it is imperative that I use an approach that addresses both scholarly communication and the institution's digital archival material." (NP respondent at a small baccalaureate college in New England)
- "Provide concise examples and talking points of benefits and successes from IRs for use to gain campus and administrative support." (NP respondent at a mid-sized public master's university in a Central Plains state)
- "Testimonials that cut to the heart of what each size institution can gain. ..." (NP respondent at a small private church-affiliated liberal arts college in a Great Lakes state)
- "Help publicize to the great diversity of academic community members the value that IRs have added to all types of institutions that have implemented them." (PO respondent at a small master's university in the Pacific Northwest)

2. Will top-rated IR benefits someday pertain to derailing the current publishing model? (See Appendixes F8.4 and F9, Table 7.1 and Subchapter 7.1, Table 7.3, and Table 7.4 and Subchapter 7.4.)

MIRACLE Project investigators included three benefits that figure prominently in discussions of the ability of IRs to derail the current publishing model (see Table 7.1 and Subchapter 7.1, and Table 7.4 and Subchapter 7.4). Although census respondents rated these three benefits in the middle of the list of 16 benefits, they were generally positive about them. Will these benefits be the ones to rise above the others in the years to come? If the future brings significant changes to today's publishing model, to what extent will IRs be responsible for the changes?

The low rate of contributions to IRs could be attributed to researchers' reluctance to upset the delicate balance between themselves and publishers (see Table 7.3). While some call for restraint, inviting the various stakeholders to partner in discussion and nego-

tiation (Lynch 1992; Drabenstott 1994; Borgman 2006), others have become activists, urging scholars to license their publications to the public domain (Creative Commons 2006), rating publishers on their self-archiving policies (RoMEO SHERPA 2006), and serving as vocal advocates of the open-access movement, e.g., Harnad 2006; Suber 1996–2006.

As more academics become aware of the evidence in favor of higher citation rates for articles published in open-access publications (see Appendix F8.4), will they be more likely to seek open-access publishers for publishing their work? What differences do scholars and scientists notice when they publish in open-access publications? Are publishers noticing changes in their relationships with authors, members of professional societies and editorial boards, and reviewers? Although others have asked these questions in the past, it has been only recently that the infrastructure for self-archiving has been in place to challenge publishers and the stronghold they have had on scholarly publishing for so long.

3. Will IRs coexist alongside subject- and discipline-oriented repositories? Or, after the dust on open access settles, will one repository type be left standing? (See Subchapter 6.5.)

The physics discipline with its arXiv subject-oriented repository is exemplary in terms of building and maintaining a successful digital repository that is embraced by the discipline as a whole (Pinfield 2001). Physicists are expected to contribute to arXiv—their standing in the field depends on it.

No other digital repository has made such deep inroads into a discipline or been met with such widespread acceptance as the physics-based arXiv digital repository. For example, Borgman (2006) cites contribution rates for subject-oriented repositories struggling to reach 15% and mentions in passing PubMed Central, where the contribution rate is a disappointing 3.8%.

The IRs that academic institutions support report low contributions rates in this very report (see Subchapter 6.5), in all previous surveys (Lynch and Lippincott 2005; Shearer 2004; Bailey et al. 2006), and in a long list of articles focusing on contributions (e.g., Foster and Gibbons 2005; Jenkins, Breakstone, and Hixson 2005; Chan, Kwok, and Yip 2005; Bell, Foster, and Gibbons 2005). Considering such low contribution rates, does it make sense for IRs and subject-oriented repositories to compete for contributors? Should one repository type yield to the other? One could imagine building functionality into the latter that automatically pops links to newly deposited material into the former. Is such double-posting necessary? Will it serve only to confuse end users searching for their topics of interest?

In the future, we might expect professional societies to enter the competition, wanting to be *the* digital repository of record for their discipline's scholarly and scientific production. At a certain point, a shakedown will occur: some repositories will merge, and others will disappear. Who will remain standing in the digital repository business remains to be seen.

4. What is the likelihood of making scholarly and scientific data and production mandatory in digital repositories? Who would police compliance with such a policy? (See Subchapter 7.3 and Table 7.3.)

MIRACLE Project questionnaires asked PO, PPT, and IMP respondents what was likely to inhibit their ability to deploy a successful IR. It listed a dozen potentially inhibiting factors and asked them to rate each factor. The three respondent types gave listed factors different ratings depending on their stage in the IR effort. PO respondents rate factors pertaining to deploying an operational IR highest, e.g., supporting the costs of an operational IR and competing with other priorities. PPT respondents are most concerned with building their IR's database with quality content, e.g., encouraging faculty to contribute to the IR and contributors' lack of knowledge of the benefits of IRs. IMP respondents, who know firsthand how difficult it is to recruit contributors and contributions to the IR, rank the "absence of campus-wide mandates regarding mandatory contribution of certain material types" as the top factor.

Policing compliance may be easier to do in IRs than in subject-oriented repositories because of promotion, tenure, and merit-increase reviews that many academic units periodically require of their research and teaching staff. Staff would be expected to link their publications to full-text sources in digital repositories. Whether they deposit in repositories the data files that they create for their research depends on their activity in humanities, social science, and science disciplines because of the different ways in which scholars create new knowledge across the disciplines (Borgman 2006).

To our knowledge, Queensland University of Technology is the only academic institution that has adopted a mandatory deposit policy for all staff members (Queensland University of Technology 2006). The Queensland mandate is undoubtedly being monitored with great interest by university administrators who are being asked to support IRs in both rhetoric and funds, by library directors who bear the brunt of the IR's support in their budgets (see Table 3.1), and by foundation, federal, and state funding officers who are eager for positive outcomes as a result the grants, contracts, and cooperative agreements they award.

5. In the absence of IR-user studies, we have enumerated several questions about IR users and uses. Answers to these questions will interest (1) decision makers at academic institutions who want to initiate an IR effort and need to convince their superiors of the benefits of IRs, (2) IR staff who recruit digital content, (3) IR-systems staff who are responsible for migrations to new versions and new systems, and (4) IR-system designers who are making improvements to existing systems and planning new ones. An IR-user study is a future MIRACLE Project activity. The list that follows provides examples of questions that could figure into such a study.

- Who uses IRs?
- For what are they searching?
- What level of perseverance do they demonstrate searching the IR?
- Why do they use IRs?
- How did they learn about the IR?
- How many times have they searched the IR in the past?
- Would they search the IR in the future? Why or why not?
- Do they understand the digital artifacts they retrieve?
- To what extent do people's searches of IRs yield relevant results?
- What exactly are relevant results?
- To what extent do people come across things of interest other than what they are looking for?
- Is a particular user type (e.g., faculty, librarians, archivists, undergraduate students) more likely to use IRs and to have success finding relevant results?
- To what purpose do people put their relevant results?
- How do they benefit from IRs?
- What help do people need before, during, and after their search of the IR?
- What improvements can be made to IR searching? To IR metadata? To IR contents? To IR controlled vocabularies?
- What other online systems do they search daily?
- What do they like about the other online systems they search that they would like to see in the IR?
- Would they recommend that their peers, colleagues, students, subordinates, etc., search the IR? Why or why not?

6. What metadata are appropriate for the wide range of artifact genres that characterize IR databases? (See Table 5.3.)

MIRACLE Project investigators were cautious about asking respondents about metadata because the closed-ended nature of questionnaires precluded achieving much depth on the topic. We will explore metadata issues in subsequent project activities. The one question that touched on metadata asked PPT and IMP respondents about IR-system features and respondents ranked features pertaining to metadata dead last (see Table 5.3). MIRACLE Project investigators will follow up on this issue and other issues pertaining to metadata in future activities.

A source of new-knowledge products and byproducts, IRs will enable scientists and scholars to find data files that pertain to their research interests, retrieve research papers that describe original and follow-up analyses using these files, and download the files for their own analyses. Unlike texts, data files are not self-describing; thus, users will benefit from metadata during both the retrieval and selection phases of the search process.

The library community that has taken the responsibility for IRs has much experience in metadata creation, but this experience pertains primarily to text-based documents. What metadata pertaining to texts are appropriate for data files? Do some metadata elements pertain to data files only? What do data-file contributors want

prospective users to know about their files that deserve to be represented in metadata? What do prospective data-file users want to know? What can librarians learn from the practice of data archivists about metadata (e.g., ICPSR 2005)? How should metadata in the IR relate to metadata in other campus information systems and library databases?

Metadata for traditional library collections have fallen short of user expectations (Markey 2007). IRs are an opportunity to start anew and to learn from the people who create the data files, the wide range of prospective users of these files, and from data archivists who have a proven track record with data files, uses, and users, so that the metadata they assign to the data files in IRs fulfill everyone's expectations.

7. In the design of IR systems and enhancement of metadata for IR content, be aware of the principle of least effort.

When staff are designing new IRs, they must keep in mind the principle of least effort. "This principle states that most researchers (even 'serious' scholars) will tend to choose easily available information sources, even when they are objectively of low quality, and, further, will tend to be satisfied with whatever can be found easily in preference to pursuing higher-quality sources whose use would require a greater expenditure of effort" (Mann 1993, 91).

IRs are likely to be a curious mix of primary, secondary, and tertiary (e.g., encyclopedias, annual reviews, yearbooks, bibliographies) sources. Humanities scholars are accustomed to searching resource-type mixtures; in fact, "recognizing something [from an archive, library, corporate records, mass media, etc.] could be a data source is a scholarly act in itself" (Borgman 2006). Researchers in science and social science disciplines are less accustomed to finding data and the research that interprets the data in one place. IR-system designers can expect people with varying levels of domain expertise—from undergraduate students to senior faculty members—to be potential users of IRs at academic institutions. A key objective for these designers should be the principle of least effort so that IRs are usable regardless of their users' domain expertise.

9.5 Chapter 9 Summary

Chapter 9 discusses the findings of the MIRACLE Project census. It begins with an examination of NP respondents, who represent the largest percentage (52%) of census respondents (see Subchapter 9.1). NP respondents come from institutions where no IR planning has been done. Dominating NPs are master's and baccalaureate institutions.

Our analysis of NP respondents reveals their great interest in IRs. They want to know how much IRs cost to plan, implement, and maintain, and what institutions comparable to their own are doing with regard to IRs (see Table 8.2 and Subchapter 9.1). None of the top-ranked reasons why NP institutions have not begun IR plan-

ning rules out their involvement with IRs at a later date (see Table 8.1). Right now, NP institutions have other things on their plate or have no resources or expertise for IR planning. Very few are totally in the dark in terms of what IRs are and whether IRs have relevance for their institutions (see Figure 8.1). Slightly less than 50% of NP respondents may start IR planning within the next 24 months (see Figure 8.2).

NP respondents would benefit from success stories about IRs from their colleagues at other-than-research-universities. If subsequent MIRACLE Project activities are biased toward the experiences of other-than-research-universities, they would focus where the need is greatest and where the gap in knowledge about IRs exists is widest.

Subchapter 9.2 enumerates census findings that verify findings from previous surveys. Subchapter 9.3 is an in-depth examination of 13 findings that are unique to the MIRACLE Project census. Subchapter 9.4 concludes the report by making observations on seven long-term issues pertaining to IRs that will continue to occupy educational institutions long after the MIRACLE Project ends. Consult the report's Executive Summary for a comprehensive treatment of MIRACLE Project census findings.

APPENDIX A

Advisory Group Members

Joseph Branin

Professor, Librarianship, and Director of Libraries
The Ohio State University, Columbus, Ohio

Michael Seadle

Director of the Institute for Library and Information Science at
Humboldt University, Berlin, Germany

Helen Tibbo

Professor, School of Library and Information Science
University of North Carolina at Chapel Hill, Chapel Hill, North
Carolina

Diane Vizine-Goetz

Consulting Research Scientist
OCLC, Dublin, Ohio

Marcia Zeng

Professor, Library and Information Science
Kent State University, Kent, Ohio

APPENDIX B

Questionnaire for No Planning (NP) Respondents

Informed Consent Form

Thank you for being willing to participate in the MIRACLE Project. Our project's objectives are to assess planning and implementation of institutional repositories (IRs) in educational institutions in the U.S. and to identify IR practices, policies, and operations. We will publish census results on the web and use them to characterize best practices and successful models of IRs.

Your participation will involve completing a web-administered questionnaire. Completing a questionnaire will take about 12 minutes. If you need to contact colleagues to answer questions, you can sign off, do your research, and sign onto the questionnaire at a later time.

At the conclusion of the questionnaire, we will ask you to volunteer your institution's name so that we can make sure that we receive completed questionnaires from all U.S. educational institutions. We will also ask you to volunteer your name in case we need clarification or you want to participate in follow-up telephone interviews and/or case studies. We will keep identifying information separate from your responses.

Later this year, results of the census will be published at the MIRACLE Project website: <http://miracle.si.umich.edu>.

Participation in the census is voluntary. You may choose not to answer some questions and you may decide to withdraw from the study at any time. There is no known risk or discomfort you will have from your participation and there is no direct benefit.

Should you have questions concerning your rights as a research participant, you should contact the Institutional Review Board, James Sayer, 540 East Liberty Street, Suite 202, Ann Arbor, MI 48104- 2210, (734) 936-0933, irbhsbs@umich.edu

For all other questions about this study please contact MIRACLE Project investigators:

Soo Young Rieh	Karen Markey	Elizabeth Yakel
Assistant Professor	Professor	Associate Professor
School of Information	School of Information	School of Information
University of Michigan	University of Michigan	University of Michigan
304 West Hall	304 West Hall	304 West Hall
1085 So. University Ave.	1085 So. University Ave.	1085 So. University Ave.
Ann Arbor, MI 48109	Ann Arbor, MI 48109	Ann Arbor, MI 48109
(734) 647-8040	(734) 763-3581	(734) 763-3569
rieh@umich.edu	ylime@umich.edu	yakel@umich.edu

I have read the information in this consent form, and I agree to participate in this study.

<input type="checkbox"/> I agree
<input type="checkbox"/> I do not agree

A. About IR Planning

1a. Please rate the importance of each reason for describing why no planning for an IR has been done at your institution to date.

	VI*	SI	SU	VU	NO	DK	NA
We want to assess IRs at other institutions generally before taking the plunge							
We want to assess IRs at institutions like our own before taking the plunge							
We are waiting to join a consortium, partnership, or group							
We do not understand or believe in the value or effectiveness of an IR							
We do not need an IR							
We will outsource IR services to another institution, consortium, partnership, or group							
We are waiting for funding to support planning							
We have no available resources to support planning							
We have no in-house expertise for planning							
Other priorities, issues, activities, etc., are more pressing than an IR							
We are not convinced that an IR would benefit our institution's learning community							
We have no support from our institution's administration							
We have no support from our library's administration							
We have no support from our institution's information technology group							
We doubt members of our institution's learning community will contribute to an IR							
Other (Please specify in question 1b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

1b. If you rated "Other" for the previous question, please specify in the box below.

2a. How important are each of the following to your institution?

	VI*	SI	SU	VU	NO	DK	NA
What other institutions generally are doing with regard to IRs							
What institutions comparable to my own are doing with regard to IRs							
What is the impetus for IR planning and implementation at other institutions generally							
What is the impetus for IR planning and implementation at institutions comparable to my own							

How to interest my institution's administration in IR planning							
How to interest my library's administration in IR planning							
How to interest an institution(s) in partnering with us on an IR							
How to interest a consortium, partnership, group, library network, etc., in IR planning							
Whether members of my institution's learning community will contribute to our IR							
Whether members of my institution's learning community will use our IR							
An IR as an accepted "best practice" in the profession							
How much it costs to plan for an IR							
How much it costs to implement an IR							
How much it costs to maintain an IR							
How much it costs to migrate to a new IR							
Other (Please specify in question 2b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

2b. If you rated "Other" for the question above, please specify in the box below.

B. Potential Contributors to an IR

3. If your institution eventually does make the decision to implement an IR, who do you think would be authorized contributors to the IR? (Choose as many as apply.)

Faculty members
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
Your institution's administrators
Your institution's press
Your institution's news service
Your institution's central computer services staff
Academic support staff
External contributors
Other (please specify)

4. If your institution eventually does make the decision to implement an IR, how easy do you think it will be to get faculty to contribute to the IR?

Very easy
Somewhat easy
Somewhat difficult
Very difficult
No opinion
Don't know
Not applicable

5. If your institution eventually does make the decision to implement an IR, how easy do you think it will be to get other members of the learning community to contribute to the IR?

Very easy
Somewhat easy
Somewhat difficult
Very difficult
No opinion
Don't know
Not applicable

C. Benefits of IRs

6a. How important do you think these anticipated benefits of IRs would be to your institution?

	VI*	SI	SU	VU	NO	DK	NA
A boost to your institution's prestige							
Better service to contributors							
Better services to your institution's learning community							
New services to learning communities beyond your institution							
Maintaining control over your institution's intellectual property							
Capturing the intellectual capital of your institution							
Contributing to the reform of the entire enterprise of scholarly communication and publishing							
A reduction in the amount of time between discovery and dissemination of research findings to scholarly communities							
An increase in citation counts to your institution's intellectual output							
Exposing your institution's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							

An increase in the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							
Providing maximal access to the results of publicly funded research							
A solution to the problem of preserving your institution's intellectual output							
An increase in your library's role as a viable partner in the research enterprise							
Reducing user dependence on your library's print collection							
Longtime preservation of your institution's digital output							
Other (Please specify in question 6b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

6b. If you rated "Other" for the question above, please specify in the box below.

D. Speculating on IR Planning in the Future

7. How likely is your institution to get involved in IR planning in the near term (the next 12 months)?

Very likely
Somewhat likely
Somewhat unlikely
Very unlikely
No opinion
Don't know
Not applicable

8. How likely is your institution to get involved in IR planning in the medium term (the next 1 to 3 years)?

Very likely
Somewhat likely
Somewhat unlikely
Very unlikely
Don't know
No opinion
Not applicable

9. What would have to happen for IR planning to begin at your institution? (Choose all that apply.)

We receive approval from our institution's administration
We receive approval from our library's administration

We receive approval from our institution’s information technology group
We receive funding from our institution’s administration
We receive funding from our library’s administration
We receive funding from our institution’s information technology group
We have evidence of successful IR demonstration projects at a comparable institution
We have evidence of successful IR demonstration projects at other institutions generally
We secure outsourcing for IR services from another institution, consortium, or group
We receive additional personnel resources to support planning
We are convinced that our institution’s learning community would contribute to it
We reassess our institution’s current priorities, issues, and activities
Other (please specify)

10a. How active do you think that the people in these positions would have to be to light the spark for IR planning at your institution?

	VA*	SA	SI	VI	NO	DK	NA
Staff at a library network, consortium, or other affiliated group							
Your institution’s president or chancellor							
Your institution’s vice president or provost							
Faculty governance, e. g., faculty senate, faculty senate assembly, etc.							
Your institution’s chief information officer							
Your institution’s archivist							
Faculty members generally							
A faculty member in particular							
Library director							
Assistant library director(s)							
Library staff member(s)							
Graduate student (s)							
Undergraduate student(s)							
Other (Please specify in question 10b below)							

* Key to abbreviations: VA=Very active, SA=Somewhat active, SI=Somewhat inactive, VI=Very inactive, NO=No opinion, DK=Don’t know, NA=Not applicable

10b. If you rated “Other” for the question above, please specify in the box below.

E. Identifying You and Your Institution

11. Please identify your position at your institution. (Choose one only.)

President or chancellor
Staff in the office of the president or chancellor
Vice president or provost
Staff in the office of the vice president or provost

Chief information officer
Staff in the office of the chief information officer
Archivist
Archives staff
Library director
Assistant director of library public services
Assistant director of library technical services
Assistant director of library information technology
Library staff
Other (please specify)

12. What is your connection to your institution's IR?
13. Please identify your institution.
14. How can the MIRACLE Project assist you regarding IRs?
15. If you would be willing to volunteer for follow-up questions via phone or email, please add your name and email address and we will contact you in the near future:

Name
Email

Thank You!

If you have questions, please message Soo Young Rieh (rieh@umich.edu) at the MIRACLE Project. Thank you for your responses.

APPENDIX C

Questionnaire for Planning Only (PO) Respondents

Informed Consent Form

(See appendix B for Informed Consent Form.)

Please answer the remaining 30 questions with the one IR in mind that offers the widest array of services to the most people and greatest number of constituencies (e.g., faculty members, students, staff, administrators, guests) in your institution's learning community. Please feel free to message Soo Young Rieh (rieh@umich.edu) with your questions or concerns.

A. Getting Started: Timelines, Funding, Benefits, and Needs Assessment

1. How long has your institution been involved in IR planning? Please enter the number of months.
2. How much longer is your institution likely to continue IR planning before making the decision whether or not to implement an IR? Please enter the number of months.
- 3a. How important are these anticipated benefits of IRs to your institution?

	VI*	SI	SU	VU	NO	DK	NA
A boost to your institution's prestige							
Better service to contributors							
Better services to your institution's learning community							
New services to learning communities beyond your institution							
Maintaining control over your institution's intellectual property							
Capturing the intellectual capital of your institution							
Contributing to the reform of the entire enterprise of scholarly communication and publishing							
A reduction in the amount of time between discovery and dissemination of research findings to scholarly communities							
An increase in citation counts to your institution's intellectual output							
Exposing your institution's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
An increase in the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							

Providing maximal access to the results of publicly funded research							
A solution to the problem of preserving your institution's intellectual output							
An increase in your library's role as a viable partner in the research enterprise							
Reducing user dependence on your library's print collection							
Longtime preservation of your institution's digital output							
Other (Please specify in question 3b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

3b. If you rated "Other" for the question above, please specify in the box below.

4. What is the likelihood that your institution will conduct a needs assessment prior to making a decision about implementing an IR?

Very likely
Somewhat likely
Somewhat unlikely
Very unlikely
No opinion
Don't know

5. Has your institution conducted a needs assessment for an IR?

Yes
No
Don't know

6a. How important are the results of the following investigative activities in terms of influencing your institution to initiate planning activities for an IR?

	VI*	SI	SU	VU	NO	DK	NA
Results of your institution's needs assessment							
Learning about successful implementations at comparable institutions							
Learning about successful implementations at a wide range of academic institutions							
Learning about available expertise and assistance from a library consortium, network, group of libraries, etc.							
An analysis of a thorough literature review of IRs							
Learning from reports of other institutions' IR planning, pilot testing IR software, and implementation activities to date							

Using other institutions' operational IRs							
Demonstrating operational IRs to my institution's decision-makers							
Demonstrating IR metadata harvesters such as OAIster and Google Scholar to my institution's decision-makers							
Waiting for a critical mass of IR implementation at comparable institutions to happen							
Waiting for a critical mass of IR implementation generally to happen							
Better digital preservation techniques							
Other (Please specify in question 6b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

6b. If you rated "Other" for the question above, please specify in the box below.

B. People Involved in IR Planning

7a. How active were people in these positions in terms of leading the charge to get involved with IRs at your institution?

	VA*	SA	SI	VI	NO	DK	NA
Staff at a library network, consortium, or other affiliated group							
Your institution's president or chancellor							
Your institution's vice president or provost							
Faculty governance, e. g., faculty senate, faculty senate assembly, etc.							
Your institution's chief information officer							
Your institution's archivist							
Faculty members generally							
A faculty member in particular							
Library director							
Assistant library director(s)							
Library staff member(s)							
Graduate student (s)							
Undergraduate student(s)							
Other (Please specify in question 7b below)							

* Key to abbreviations: VA=Very active, SA=Somewhat active, SI=Somewhat inactive, VI=Very inactive, NO=No opinion, DK=Don't know, NA=Not applicable

7b. If you rated "Other" for the question above, please specify in the box below.

8. Who is the individual leading IR planning at your institution? (Choose one only.)

A faculty members in a particular college, department, or school
Your institution's chief information officer
Your institution's archivist
Library director
Assistant library director
A library staff member
No committee or committee chair has been appointed
Other (please specify)

9. If a committee is involved with IR planning, identify the positions of other people on this committee. (Choose all that apply.)

Staff from the office of the president or chancellor
Staff from the office of the vice-president or provost
Staff from the office of the chief information officer
Staff from your institution's legal office
Your institution's chief information officer
Your institution's archivist
Library director
Assistant library director
Library staff member(s)
Archives staff
A faculty member in particular
Graduate student(s)
Undergraduate student(s)
Committee members have not yet been appointed
Other (please specify)

10. How many people are involved in IR planning?

11a. What percentage of the responsibility for an operational IR do you think should be given to various campus units? (Percentages must add up to 100%.)

	% Your institution's central administration
	% Your institution's library
	% Your institution's central computing unit
	% The office of the chief information officer
	% Your institution's archives
	% Various academic colleges, departments, and schools
	% Other (Please specify in question 11b below)

11b. If you provided a percentage for "Other" for the question above, please specify in the box below.

C. Contributors to the IR

12. If you could foretell the future, who will be authorized contributors to your institution's IR? (Choose as many as apply.)

Faculty members
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
Your institution's administrators
Your institution's press
Your institution's news service
Your institution's central computer services staff
Academic support staff
External contributors
Other (please specify)

13. Who do you think will be the major contributor to your institution's IR? (Choose one only.)

Faculty
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
University and college administrators
Computer services staff
Academic support staff
Other (please specify)

14. If your institution eventually does make the decision to implement an IR, how easy do you think it will be to get faculty to contribute to the IR?

Very easy
Somewhat easy
Somewhat difficult
Very difficult
No Opinion
Don't know
Not applicable

15. If your institution eventually does make the decision to implement an IR, how easy do you think it will be to get other members of the learning community to contribute to the IR?

Very easy
Somewhat easy
Somewhat difficult
Very difficult
No opinion
Don't know
Not applicable

16a. Why do you think members of your institution's learning community will contribute to an IR?

	VI*	SI	SU	VU	NO	DK	NA
To boost the particular scholar's prestige							
To boost your institution's prestige							
To contribute to the reform of the entire enterprise of scholarly communication and publishing							
To reduce the amount of time between discovery and dissemination of research findings to scholarly communities							
To increase citation counts to the particular scholar's oeuvre							
To increase citation counts to your institution's intellectual output							
To encourage other scholars to provide open access to their intellectual output							
To expose the particular scholar's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
To expose your institution's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
To place the burden of preservation on the IR instead of on individual faculty members							
To increase the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							
To provide maximal access to the results of publicly funded research							
To solve the problem of preserving your institution's intellectual output							
To increase the library's role as a viable partner in the research enterprise							
To reduce user dependence on your library's print collection							
Other (Please specify in question 16b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

16b. If you rated "Other" for the question above, please specify in the box below.

D. Digital Content for the IR

17a. What digital content recruitment methods do you think will be most successful at your institution?

	VS*	SS	SU	VU	NO	DK	NA
Volunteer contributions							
Publicity about the IR in campus newspapers							
Presentations by staff responsible for the IR at departmental and faculty meetings							
Personal visits by staff responsible for the IR to faculty and administrators							
Staff responsible for the IR working one-on-one with early adopters							
Word-of-mouth from early adopters to their colleagues in the faculty and staff ranks							
Publicizing the IR during reference interactions in libraries and archives							
Systematic review of faculty, staff, center, and departmental web sites for potential contributors by staff responsible for the IR							
Institution-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints, etc.							
Other (Please specify in question 17b below)							

* Key to abbreviations: VS=Very successful, SS=Somewhat successful, SU=Somewhat unsuccessful, VU=Very unsuccessful, NO=No opinion, DK=Don't know, NA=Not applicable

17b. If you rated "Other" for the question above, please specify in the box below.

18. Check all the types of digital materials you are planning to allow authorized contributors to submit to your institution's IR.

Preprints
Working papers
Books
Journals
Journal articles
Maps
Interview transcripts
Sound recordings of interview transcripts
Software
Software documentation
Video recordings of performances
Blogs

Interim and final reports to funding agencies
Raw data files that result from faculty research projects
Raw data files that result from doctoral dissertation research
Raw data files that result from master's thesis research
Raw data files that result from senior thesis research
Written papers or transcripts of conference presentations
Conference presentations (e.g., summaries, abstracts, notes, outlines, remarks, etc.)
Committee meeting agenda and minutes
Committee meeting documents, e.g., budgets, reports, memoranda
Your institution's course catalogs
Your institution's newspapers
Your institution's alumni publications
Faculty senate agendas and minutes
College, departmental, and school alumni publications
Regent, trustee, board meeting agenda and minutes
Course syllabi, class notes, handouts, outlines, assignments prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Other learning objects such as simulations, models, software demonstration files, images, video prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Doctoral dissertations
Master's theses
Senior theses
Graduate student eportfolios
Undergraduate student eportfolios
Class notes, outlines, assignments, papers, and projects prepared by graduate students
Class notes, outlines, assignments, papers, and projects prepared by undergraduate students
Other (please specify)

19. Who will be responsible for managing the IR's intellectual property rights? (Choose all that apply.)

Contributors' academic or service unit
One chosen academic unit
One chosen service unit
IR staff
Library staff
Archives staff
Staff from the office of the chief information office
A company that our IR is outsourcing
Other (please specify)

E. Speculating on Your IR’s Future

20a. How likely are each of the following to be your next steps on the road to IR implementation as a direct result of your IR planning?

	VL*	SL	SU	VU	NO	DK	NA
Your institution terminates its investigation of IRs							
Your institution seeks a partner institution(s) to share in an IR							
Your institution seeks funding for the next step of investigation of IRs							
Your institution supports implementation of an IR software package							
Your institution widens the scope of its investigation into IRs							
Your institution waits for a consortium, network, group, etc., to implement an IR							
Other (Please specify in question 20b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don’t know, NA=Not applicable

20b. If you rated “Other” for the question above, please specify in the box below.

21. What is the likelihood that your institution will pilot test one or more IR software package(s) prior to making a decision about implementing an IR?

Very likely
Somewhat likely
Somewhat unlikely
Very unlikely
No opinion
Don’t know

22a. What is the status of these IR policies?

	NP*	D	I	DK	NA
Determining what is acceptable content					
Defining collections					
Determining who is authorized to make contributions to the IR					
Restricting access to IR content					
Acceptable file formats					
Identifying metadata formats and authorized metadata creators					
Charging for IR services					
Formulating a privacy policy for registered IR system users					
Licensing IR content					
Updating IR content					
Withdrawing IR content					

Providing access management services					
Preserving IR content					
Revising IR policies in the future					
Authorizing external contributors					
Intellectual property					
Other (Please specify in question 22b below)					

* Key to abbreviations: NP=No policy; D=Drafted; I=Implemented; DK=Don't know, NA=Not applicable

22b. If you rated "Other" for the question above, please specify in the box below.

23a. To what extent do you think the following are likely to inhibit your ability to deploy a successful IR?

	VL*	SL	SU	VU	NO	DK	NA
Making members of your institution's learning community aware of the IR							
Contributors' lack of knowledge about how they can benefit from IRs							
Encouraging faculty to submit digital content to the IR							
Convincing faculty that the IR will not adversely affect the current publishing model							
Absence of campus-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints, etc.							
Contributors' concerns about the difficulty using the IR system to contribute digital content to the IR							
Inability of contributors to formulate quality metadata							
Contributors' concerns about intellectual property rights for digital materials							
Inadequacy of the IR system's digital preservation capabilities							
Difficulties in long-term preservation of digital files							
Lack of on-campus technical expertise in IR systems							
Supporting all ongoing costs of an operational IR							
Competing for resources with other priorities, projects, and initiatives							
Other (Please specify in question 23b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

23b. If you rated "Other" for the question above, please specify in the box below.

24. To what extent will an IR affect your institution's ability to build relationships between the IR and other on-campus repositories (e.g., archives, student services, library systems, digital asset management systems, electronic course management systems, digital libraries)?

A big positive effect
A moderate positive effect
No effect
A moderate negative effect
A big negative effect
A combination of positive and negative effects
Don't know
No opinion
Not applicable
Other (please specify)

25a. How likely is it that funding for your institution's implementation of an IR will come from these sources?

	VL*	SL	SU	VU	NO	DK	NA
Special initiative supported by your institution's central administration							
Special initiative supported by your institution's library							
Special initiative supported by your institution's central computer services							
Special initiative supported by your institution's archives							
Special initiative supported by academic colleges, departments, and schools							
Regular budget line item for your institution's central administration							
Regular budget line item for your institution's library							
Regular budget line item for your institution's central computer services							
Regular budget line item for your institution's archives							
Regular budget line item for academic colleges, departments, and schools							
Costs absorbed in routine operating costs of your institution's central administration							
Costs absorbed in routine operating costs of your institution's library							
Costs absorbed in routine operating costs of your institution's central computer services							
Costs absorbed in routine operating costs of your institution's archives							
Costs absorbed in routine operating costs of your institution's academic colleges, departments, and schools							
Grant awarded by an external source							

Grant awarded by an internal source							
Other (Please specify in question 25b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

25b. If you rated "Other" for the question above, please specify in the box below.

F. Identifying You and Your Institution

26. Please identify your position at your institution. (Choose one only.)

President or chancellor
Staff in the office of the president or chancellor
Vice president or provost
Staff in the office of the vice president or provost
Chief information officer
Staff in the office of the chief information officer
Archivist
Archives staff
Library director
Assistant director of library public services
Assistant director of library technical services
Assistant director of library information technology
Library staff
Other (please specify)

27. What is your connection to your institution's IR?

28. Please identify your institution.

G. Follow-up information

29. How can the MIRACLE Project assist you regarding IRs?

30. If you would be willing to volunteer for follow-up questions via phone or email, please add your name and email address and we will contact you in the near future:

Name
Email

Thank You! If you have questions, please message Soo Young Rieh (rieh@umich.edu) at the MIRACLE Project. Thank you for your responses.

APPENDIX D

Questionnaire for Planning and Pilot Testing (PPT) Respondents

Informed Consent Form

(See appendix B for Informed Consent Form.)

A. Number of IRs

- How many institutional repositories (IRs)—general IRs, special-purpose IRs, and IRs in the pilot-testing phase—are available or will be available to members of your institution’s learning community in the near future?

1
2
3
4
5 or more

Please answer the remaining 39 questions with the one IR in mind that offers the widest array of services to the most people and greatest number of constituencies (e.g., faculty members, students, staff, administrators, guests) in your institution’s learning community. Please feel free to message Soo Young Rieh (rieh@umich.edu) with your questions or concerns.

B. Getting Started: Timelines, Funding, Benefits, and Needs Assessment

- How long has your institution been involved with IR planning and pilot testing? Please enter the number of months.
- How much longer is your institution likely to continue IR planning and pilot testing before making the decision whether or not to implement an IR? Please enter the number of months.
- 4a. How important are these anticipated benefits of IRs to your institution?

	VI*	SI	SU	VU	NO	DK	NA
A boost to your institution’s prestige							
Better service to contributors							
Better services to your institution’s learning community							
New services to learning communities beyond your institution							
Maintaining control over your institution’s intellectual property							
Capturing the intellectual capital of your institution							

Contributing to the reform of the entire enterprise of scholarly communication and publishing							
A reduction in the amount of time between discovery and dissemination of research findings to scholarly communities							
An increase in citation counts to your institution's intellectual output							
Exposing your institution's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
An increase in the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							
Providing maximal access to the results of publicly funded research							
A solution to the problem of preserving your institution's intellectual output							
An increase in your library's role as a viable partner in the research enterprise							
Reducing user dependence on your library's print collection							
Longtime preservation of your institution's digital output							
Other (Please specify in question 4b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

- 4b. If you rated "Other" for the question above, please specify in the box below.
5. What is the likelihood that your institution will conduct a needs assessment prior to making a decision about implementing an IR?

Very likely
Somewhat likely
Somewhat unlikely
Very unlikely
No opinion
Don't know

6. Has your institution conducted a needs assessment for an IR?

Yes
No
Don't know

- 7a. How important are the results of the following investigative activities in terms of influencing your institution to initiate planning and pilot testing activities for an IR?

	VI*	SI	SU	VU	NO	DK	NA
Results of your institution's needs assessment							
Learning about successful implementations at comparable institutions							
Learning about successful implementations at a wide range of academic institutions							
Learning about available expertise and assistance from a library consortium, network, group of libraries, etc.							
An analysis of a thorough literature review of IRs							
Learning from reports of other institutions' IR planning, pilot testing IR software, and implementation activities to date							
Using other institutions' operational IRs							
Demonstrating operational IRs to my institution's decision-makers							
Demonstrating IR metadata harvesters such as OAIster and Google Scholar to my institution's decision-makers							
Waiting for a critical mass of IR implementation at comparable institutions to happen							
Waiting for a critical mass of IR implementation generally to happen							
Better digital preservation techniques							
Other (Please specify in question 7b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

7b. If you rated "Other" for the question above, please specify in the box below.

C. People Involved in IR Planning and Pilot Testing

8a. How active were people in the following positions in terms of leading the charge to get involved with IRs at your institution?

	VA*	SA	SI	VI	NO	DK	NA
Staff at a library network, consortium, or other affiliated group							
Your institution's president or chancellor							
Your institution's vice president or provost							
Faculty governance, e. g., faculty senate, faculty senate assembly, etc.							
Your institution's chief information officer							
Your institution's archivist							
Faculty members generally							
A faculty member in particular							

Library director							
Assistant library director(s)							
Library staff member(s)							
Graduate student (s)							
Undergraduate student(s)							
Other (Please specify in question 8b below)							

* Key to abbreviations: VA=Very active, SA=Somewhat active, SI=Somewhat inactive, VI=Very inactive, NO=No opinion, DK=Don't know, NA=Not applicable

8b. If you rated "Other" for the question above, please specify in the box below.

9. Who is the individual leading IR planning and pilot testing at your institution? (Choose one only.)

A faculty member in a particular college, department, or school
Your institution's chief information officer
Your institution's archivist
Library director
Assistant library director
A library staff member
No committee or committee chair has been appointed
Other (please specify)

10. If a committee is involved in IR planning and pilot testing, identify the positions of the other people on this committee. (Choose as many as apply.)

Staff from the office of the president or chancellor
Staff from the office of the vice-president or provost
Staff from the office of the chief information officer
Staff from your institution's legal office
Your institution's chief information officer
Your institution's archivist
Library director
Assistant library director
Library staff member(s)
Archives staff
A faculty member in particular
Graduate student(s)
Undergraduate student(s)
Committee members have not yet been appointed
Other (please specify)

11. How many people are involved in IR planning and pilot testing?

12a. What percentage of the responsibility for an operational IR do you think should be given to various campus units? (Percentages must add up to 100%.)

	% Your institution's central administration
	% Your institution's library
	% Your institution's central computing unit
	% The office of the chief information officer
	% Your institution's archives
	% Various academic colleges, departments, and schools
	% Other (Please specify in question 12b below)

12b. If you provided a percentage for 'Other' for the question above, please specify in the box below.

D. Contributors to the IR

13. If you could foretell the future, who will be authorized contributors to your institution's IR? (Choose as many as apply.)

Faculty members
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
Your institution's administrators
Your institution's press
Your institution's news service
Your institution's central computer services staff
Academic support staff
External contributors
Other (please specify)

14. Who do you think will be the major contributor to your institution's IR? (Choose one only.)

Faculty
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
University and college administrators
Computer services staff
Academic support staff
Other (please specify)

15a. Why do you think members of your institution's learning community will contribute to an IR?

	VI*	SI	SU	VU	NO	DK	NA
To boost the particular scholar's prestige							
To boost your institution's prestige							
To contribute to the reform of the entire enterprise of scholarly communication and publishing							
To reduce the amount of time between discovery and dissemination of research findings to scholarly communities							
To increase citation counts to the particular scholar's oeuvre							
To increase citation counts to your institution's intellectual output							
To encourage other scholars to provide open access to their intellectual output							
To expose the particular scholar's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
To expose your institution's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
To place the burden of preservation on the IR instead of on individual faculty members							
To increase the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							
To provide maximal access to the results of publicly funded research							
To solve the problem of preserving your institution's intellectual output							
To increase the library's role as a viable partner in the research enterprise							
To reduce user dependence on your library's print collection							
Other (Please specify in question 15b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

15b. If you rated "Other" for the question above, please specify in the box below.

16a. What digital content recruitment methods do you think will be most successful at your institution?

	VS*	SS	SU	VU	NO	DK	NA
Volunteer contributions							
Publicity about the IR in campus newspapers							
Presentations by staff responsible for the IR at departmental and faculty meetings							

Personal visits by staff responsible for the IR to faculty and administrators							
Staff responsible for the IR working one-on-one with early adopters							
Word-of-mouth from early adopters to their colleagues in the faculty and staff ranks							
Publicizing the IR during reference interactions in libraries and archives							
Systematic review of faculty, staff, center, and departmental web sites for potential contributors by staff responsible for the IR							
Institution-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints, etc.							
Other (Please specify in question 16b below)							

* Key to abbreviations: VS=Very successful, SS=Somewhat successful, SU=Somewhat unsuccessful, VU=Very unsuccessful, NO=No opinion, DK=Don't know, NA=Not applicable

16b. If you rated "Other" for the question above, please specify in the box below.

E. Pilot Testing IR Software Packages

17. What IR software packages are you pilot testing? (Choose all that apply.)

ARNO
bePress
CDSWare
ContentDM
DigiTool (Ex Libris)
DiVA
Documentum
Dpubs
DSpace
Fedora
GNU Eprints
Greenstone
HarvestRoad Hive
Innovative Interfaces
i-TOR
Luna
myCORE
OPUS
Sunsite
Virginia Tech ETD software
None
Other (please specify)

18. What interoperability standards do you want your IR to support? (Choose all that apply.)

IR supports OAI-MPH
IR is OpenURL compliant
IR materials use persistent identifiers
Our institution's federated searching includes the IR
Other (please specify)

19a. What are the most important benefits of having pilot tested one or more IRs?

	VI*	SI	SU	VU	NO	DK	NA
Giving demonstrations to people involved in the IR implementation decision							
Giving demonstrations to an institution (s) interested in partnering with us to encourage them in IR implementation							
Gauging the interest of potential contributors to the IR							
Gauging the interest of potential IR-system users							
Identifying the strengths and shortcomings of available IR software							
Estimating costs for the technical implementation of an operational IR							
Developing the requisite technical expertise for IR implementation							
Identifying first adopters of an IR at your institution							
Control over your institution's intellectual output							
Preservation of your institution's intellectual output							
Other (Please specify in question 19b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

19b. If you rated "Other" for the question above, please specify in the box below.

20a. Based on your pilot testing of IR software packages, how would you rate IR systems generally with regard to these capabilities?

	VA*	SA	SI	VI	NO	DK	NA
Technical support							
Technical documentation							
Adherence to open access standards							
Scalability = System growth and enhancement							
Customization							
Extensibility = Access to other campus systems and data							
Supported file formats							
User authentication							

Formulating metadata for digital documents							
Browsing, searching, and retrieving digital content							
End-user interface generally							
Controlled vocabulary searching							
Authority control							
Digital preservation							
Other (Please specify in question 20b below)							

* Key to abbreviations: VA=Very adequate, SA=Somewhat adequate, SI=Somewhat inadequate, VI=Very inadequate, NO=No opinion, DK=Don't know, NA=Not applicable

- 20b. If you rated "Other" for the question above, please specify in the box below.
21. If your pilot testing with IRs involves early adopters of IR technology, from what academic colleges, departments, schools, and service units will they come? (Choose all that apply.)

Your institution's library
Your institution's central computing unit
Your institution's archives
A particular academic college, department, or school
A particular service unit
Don't know
Not applicable
Other (please specify)

F. Digital Content for the IR

22. What digital documents make up your IR's collections in its present pilot-testing phase? (Mark all that apply.)

Preprints
Working papers
Books
Journals
Journal articles
Maps
Interview transcripts
Sound recordings of interview transcripts
Software
Software documentation
Video recordings of performances
Blogs
Interim and final reports to funding agencies
Raw data files that result from faculty research projects
Raw data files that result from doctoral dissertation research
Raw data files that result from master's thesis research

Raw data files that result from senior thesis research
Written papers or transcripts of conference presentations
Conference presentations (e.g., summaries, abstracts, notes, outlines, remarks, etc.)
Committee meeting agenda and minutes
Committee meeting documents, e.g., budgets, reports, memoranda
Your institution's course catalogs
Your institution's newspapers
Your institution's alumni publications
Faculty senate agendas and minutes
College, departmental, and school alumni publications
Regent, trustee, board meeting agenda and minutes
Course syllabi, class notes, handouts, outlines, assignments prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Other learning objects such as simulations, models, software demonstration files, images, video prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Doctoral dissertations
Master's theses
Senior theses
Graduate student eportfolios
Undergraduate student eportfolios
Class notes, outlines, assignments, papers, and projects prepared by graduate students
Class notes, outlines, assignments, papers, and projects prepared by undergraduate students
Other (please specify)

23. Estimate the total number of digital documents that are published or in process in the IR that you are pilot testing.
- 24a. Estimate the number of digital documents that make up your IR's collections in its pilot-testing phase. (Write in the amount or write in DK for Don't Know or NA for Not Applicable.)

Preprints
Working papers
Books
Journals
Journal articles
Maps
Interview transcripts
Sound recordings of interview transcripts
Software
Software documentation
Video recordings of performances
Blogs
Interim and final reports to funding agencies
Raw data files that result from faculty research projects

Raw data files that result from doctoral dissertation research
Raw data files that result from master's thesis research
Raw data files that result from senior thesis research
Written papers or transcripts of conference presentations
Conference presentations (e.g., summaries, abstracts, notes, outlines, remarks, etc.)
Committee meeting agenda and minutes
Committee meeting documents, e.g., budgets, reports, memoranda
Your institution's course catalogs
Your institution's newspapers
Your institution's alumni publications
Faculty senate agendas and minutes
College, departmental, and school alumni publications
Regent, trustee, board meeting agendas and minutes
Course syllabi, class notes, handouts, outlines, assignments prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Other learning objects such as simulations, models, software demonstration files, images, video prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Doctoral dissertations
Master's theses
Senior theses
Graduate student eportfolios
Undergraduate student eportfolios
Class notes, outlines, assignments, papers, and projects prepared by graduate students
Class notes, outlines, assignments, papers, and projects prepared by undergraduate students
Other (Please specify type of digital document in question 24b below)

24b. If you entered an estimate for "Other" in the previous question, please specify in the box below.

25. When you formally implement an IR, do you intend to add the same kinds of digital content into the system?

Yes, the same kinds
Yes, the same and other kinds of content
No
Don't know
Not applicable
Maybe. Please explain:

26a. What file formats have you guaranteed contributors that you will preserve in perpetuity?

	Guaranteed	DK*	NO	NA
Plain Text UTF-8 (Unicode)				
Plain Text ANSI X3.4/ECMA-6/US-ASCII (7-bit)				
Plain Text ISO 8859-x (8-bit)				
Plain Text (all other encodings, including, but not limited to ISO 646 national variants)				
Rich text				
XML				
TeX				
LaTeX				
Postscript				
PDF				
PDF/A				
Microsoft Word				
Microsoft Excel				
Microsoft PowerPoint				
TIFF				
GIF				
JPEG				
PNG				
BMP				
Photo CD				
Photoshop				
AIFF				
Audio/Basic				
MPEG audio				
AAC_M4A				
Real Audio				
Windows Media Audio				
Wave				
AVI				
MPEG-1				
MPEG-2				
MPEG-4				
Windows Media Video				
Quicktime				
Other (Please specify file format in question 26b below)				

* Key to abbreviations: NO=No opinion, DK=Don't know, NA=Not applicable

26b. If you rated "Other" for the question above, please specify in the box below.

27. Who will be responsible for managing the IR's intellectual property rights? (Choose all that apply.)

Contributors' academic or service unit
One chosen academic unit
One chosen service unit
IR staff
Library staff
Archives staff
Staff from the office of the chief information office
A company that our IR is outsourcing
Other (please specify)

G. Speculating on Your IR's Future

28a. How likely are each of the following to be your next steps on the road to IR implementation as a direct result of your IR planning and pilot testing?

	VL*	SL	SU	VU	NO	DK	NA
Your institution terminates its investigation of IRs							
Your institution seeks a partner institution(s) to share in an IR							
Your institution seeks funding for the next step of investigation of IRs							
Your institution supports implementation of an IR software package							
Your institution widens the scope of its investigation into IRs							
Your institution waits for a consortium, network, group, etc., to implement an IR							
Other (Please specify in question 28b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

28b. If you rated "Other" for the question above, please specify in the box below.

29a. What is the status of these IR policies?

	NP*	D	I	DK	NA
Determining what is acceptable content					
Defining collections					
Determining who is authorized to make contributions to the IR					
Restricting access to IR content					
Acceptable file formats					
Identifying metadata formats and authorized metadata creators					
Charging for IR services					

Formulating a privacy policy for registered IR system users					
Licensing IR content					
Updating IR content					
Withdrawing IR content					
Providing access management services					
Preserving IR content					
Revising IR policies in the future					
Authorizing external contributors					
Intellectual property					
Other (Please specify in question 29b below)					

* Key to abbreviations: NP=No policy; D=Drafted; I=Implemented; DK=Don't know, NA=Not applicable

29b. If you rated "Other" for the question above, please specify in the box below.

30a. To what extent do you think the following are likely to inhibit your ability to deploy a successful IR?

	VL*	SL	SU	VU	NO	DK	NA
Making members of your institution's learning community aware of the IR							
Contributors' lack of knowledge about how they can benefit from IRs							
Encouraging faculty to submit digital content to the IR							
Convincing faculty that the IR will not adversely affect the current publishing model							
Absence of campus-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints, etc.							
Contributors' concerns about the difficulty using the IR system to contribute digital content to the IR							
Inability of contributors to formulate quality metadata							
Contributors' concerns about intellectual property rights for digital materials							
Inadequacy of the IR system's digital preservation capabilities							
Difficulties in long-term preservation of digital files							
Lack of on-campus technical expertise in IR systems							
Supporting all ongoing costs of an operational IR							
Competing for resources with other priorities, projects, and initiatives							
Other (Please specify in question 30b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

- 30b. If you rated "Other" for the question above, please specify in the box below.
31. To what extent will an IR affect your institution's ability to build relationships between the IR and other on-campus repositories (e.g., archives, student services, library systems, digital asset management systems, electronic course management systems, digital libraries)?

A big positive effect
A moderate positive effect
No effect
A moderate negative effect
A big negative effect
A combination of positive and negative effects
Don't know
No opinion
Not applicable
Other (please specify)

- 32a. How likely is it that funding for your institution's implementation of an IR will come from these sources?

	VL*	SL	SU	VU	NO	DK	NA
Special initiative supported by your institution's central administration							
Special initiative supported by your institution's library							
Special initiative supported by your institution's central computer services							
Special initiative supported by your institution's archives							
Special initiative supported by academic colleges, departments, and schools							
Regular budget line item for your institution's central administration							
Regular budget line item for your institution's library							
Regular budget line item for your institution's central computer services							
Regular budget line item for your institution's archives							
Regular budget line item for academic colleges, departments, and schools							
Costs absorbed in routine operating costs of your institution's central administration							
Costs absorbed in routine operating costs of your institution's library							
Costs absorbed in routine operating costs of your institution's central computer services							

Costs absorbed in routine operating costs of your institution's archives							
Costs absorbed in routine operating costs of your institution's academic colleges, departments, and schools							
Grant awarded by an external source							
Grant awarded by an internal source							
Other (Please specify in question 32b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

32b. If you rated "Other" for the question above, please specify in the box below.

33a. What percentage of your IR's annual budget is allocated to these categories? (Percentages must add up to 100%.)

	% Staff (including benefits)
	% Hardware acquisition
	% Hardware maintenance
	% Software acquisition
	% Software maintenance and updates
	% System backup
	% Vendor fees (for IRs hosted by an external vendor)
	% Other (Please specify in 33b question below)

33b. If you provided a percentage for "Other" for the question above, please specify in the box below.

H. Identifying You and Your Institution

34. Please identify your position at your institution. (Choose one only.)

President or chancellor
Staff in the office of the president or chancellor
Vice president or provost
Staff in the office of the vice president or provost
Chief information officer
Staff in the office of the chief information officer
Archivist
Archives staff
Library director
Assistant director of library public services
Assistant director of library technical services
Assistant director of library information technology
Library staff
Other (please specify)

35. What is your connection to your institution's IR?

36. Please identify your institution.
37. If the IRs your institution is pilot testing are available to the general public, please give their web address(es).

I. Follow-up information

38. How can the MIRACLE Project assist you regarding IRs?
39. If you would be willing to volunteer for follow-up questions via phone or email, please add your name and email address and we will contact you in the near future:

Name
Email

Thank you! If you have questions, please message Soo Young Rieh (rieh@umich.edu) at the MIRACLE Project. Thank you for your responses.

APPENDIX E

Questionnaire for Implementation (IMP) Respondents

Informed Consent Form

(See appendix B for Informed Consent Form.)

A. Number of IRs

1. How many institutional repositories (IRs)—general IRs, special-purpose IRs, and IRs in the pilot-testing phase—are available or will be available to members of your institution’s learning community in the near future?

1
2
3
4
5 or more

B. Specific IR implementation

Please answer the remaining 42 questions with the one IR in mind that offers the widest array of services to the most people and greatest number of constituencies (e.g., faculty members, students, staff, administrators, guests) in your institution’s learning community. Please feel free to message Soo Young Rieh (rieh@umich.edu) with your questions or concerns.

C. Timelines and Funding

2. How long has your institution been involved with IRs (everything from planning, pilot testing IR systems, to system implementation)? Please enter the number of months.
3. How long has your IR been operational, that is, available to authorized users for submission and searching of digital content? Please enter the number of months.

D. Needs Assessment

4. Did your institution conduct a needs assessment prior to implementing an IR?

Yes
No
Don’t know

5a. How important were the results of the needs assessment for:

	VI*	SI	SU	VU	NO	DK	NA
Identifying first adopters of an IR							
Identifying especially active contributors to the IR							
Formulating IR policies							
Making the decision to implement an IR							
Increasing faculty awareness of the IR							
Recruiting digital content for the IR							
Streamlining IR planning and implementation							
Choosing an IR software package							
Scheduling the rollout of various IR services							
Identifying new services to build onto the IR							
Identifying preservation techniques							
Other (Please specify in question 5b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

5b. If you rated "Other" for the question above, please specify in the box below.

E. Influences on IR Implementation Decision

6a. How important were the results of the following investigative activities in terms of influencing your institution's decision about implementing an IR?

	VI*	SI	SU	VU	NO	DK	NA
Results of your institution's needs assessment							
Learning about successful implementations at comparable institutions							
Learning about successful implementations at a wide range of academic institutions							
Learning about available expertise and assistance from a library consortium, network, group of libraries, etc.							
An analysis of a thorough literature review of IRs							
Learning from reports of other institutions' IR planning, pilot testing IR software, and implementation activities to date							
Using other institutions' operational IRs							
Demonstrating operational IRs to my institution's decision-makers							
Demonstrating IR metadata harvesters such as OAIster and Google Scholar to my institution's decision-makers							
Waiting for a critical mass of IR implementation at comparable institutions to happen							

Waiting for a critical mass of IR implementation generally to happen							
Identifying better digital preservation techniques							
Other (Please specify in question 6b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

6b. If you rated "Other" for the question above, please specify in the box below.

F. Benefits of IRs

7a. At the beginning of IR planning at your institution, how important did you think these anticipated benefits of IRs would be to your institution?

	VI*	SI	SU	VU	NO	DK	NA
A boost to your institution's prestige							
Better service to contributors							
Better services to your institution's learning community							
New services to learning communities beyond your institution							
Maintaining control over your institution's intellectual property							
Capturing the intellectual capital of your institution							
Contributing to the reform of the entire enterprise of scholarly communication and publishing							
A reduction in the amount of time between discovery and dissemination of research findings to scholarly communities							
An increase in citation counts to your institution's intellectual output							
Exposing your institution's intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
An increase in the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							
Providing maximal access to the results of publicly funded research							
A solution to the problem of preserving your institution's intellectual output							
An increase in your library's role as a viable partner in the research enterprise							
Reducing user dependence on your library's print collection							

Providing maximal access to the results of publicly funded research									
A solution to the problem of preserving your institution's intellectual output									
An increase in your library's role as a viable partner in the research enterprise									
Reducing user dependence on your library's print collection									
Longtime preservation of your institution's digital output									
Other (Please specify in question 8b below)									

* Key to abbreviations: VMMI=Very much more important, SMI=Somewhat more important, NC=No change in importance; SLI=Somewhat less important, VMLI=Very much less important, NO=No opinion, DK=Don't know, NA=Not applicable

8b. If you rated "Other" for the question above, please specify in the box below.

G. People involved in the IR effort

9a. How active were people in the following positions in terms of leading the charge to get involved with IRs at your institution?

	VA*	SA	SI	VI	NO	DK	NA
Staff at a library network, consortium, or other affiliated group							
Your institution's president or chancellor							
Your institution's vice president or provost							
Faculty governance, e. g., faculty senate, faculty senate assembly, etc.							
Your institution's chief information officer							
Your institution's archivist							
Faculty members generally							
A faculty member in particular							
Library director							
Assistant library director(s)							
Library staff member(s)							
Graduate student (s)							
Undergraduate student(s)							
Other (Please specify in question 9b below)							

* Key to abbreviations: VA=Very active, SA=Somewhat active, SI=Somewhat inactive, VI=Very inactive, NO=No opinion, DK=Don't know, NA=Not applicable

9b. If you rated "Other" for the question above, please specify in the box below.

10. Who is the individual leading IR implementation at your institution? (Choose one only.)

A faculty member in a particular college, department, or school
Your institution's chief information officer
Your institution's archivist
Library director
Assistant library director
A library staff member
No committee or committee chair has been appointed
Other (please specify)

11. If a committee is involved with IR implementation, identify the positions of the other people on this committee. (Please check all that apply.)

Staff from the office of the president or chancellor
Staff from the office of the vice-president or provost
Staff from the office of the chief information officer
Staff from your institution's legal office
Your institution's chief information officer
Your institution's archivist
Library director
Assistant library director
Library staff member(s)
Archives staff
A faculty member in particular
Graduate student(s)
Undergraduate student(s)
Committee members have not yet been appointed
Other (please specify)

12. How many people are involved in your institution's IR implementation?

H. IR Responsibility

13a. What percentage of the responsibility for an operational IR has been given to various campus units? (Percentages must add up to 100%.)

	% Your institution's central administration
	% Your institution's library
	% Your institution's central computing unit
	% The office of the chief information officer
	% Your institution's archives
	% Various academic colleges, departments, and schools
	% Other (Please specify in question 13b below)

13b. If you provided a percentage for "Other" for the question above, please specify in the box below.

I. Contributions to the IR

14. Who are authorized contributors to your institution’s IR? (Choose as many as apply.)

Faculty members
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
Your institution’s administrators
Your institution’s press
Your institution’s news service
Your institution’s central computer services staff
Academic support staff
External contributors
Other (please specify)

15. Who is the major contributor to your institution’s IR? (Choose one only.)

Faculty
Graduate students
Undergraduate students
Research scientists
Librarians
Archivists
University and college administrators
Computer services staff
Academic support staff
Other (please specify)

16a. When planning for an IR, what did you think would be the most important reasons why members of your institution’s learning community would contribute to the IR?

	VI*	SI	SU	VU	NO	DK	NA
To boost the particular scholar’s prestige							
To boost your institution’s prestige							
To contribute to the reform of the entire enterprise of scholarly communication and publishing							
To reduce the amount of time between discovery and dissemination of research findings to scholarly communities							
To increase citation counts to the particular scholar’s oeuvre							
To increase citation counts to your institution’s intellectual output							

To encourage other scholars to provide open access to their intellectual output							
To expose the particular scholar’s intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
To expose your institution’s intellectual output to researchers in North America and around the world who would not otherwise have access to it through traditional channels							
To place the burden of preservation on the IR instead of on individual faculty members							
To increase the accessibility to knowledge assets such as numeric, video, audio, and multimedia datasets							
To provide maximal access to the results of publicly funded research							
To solve the problem of preserving your institution’s intellectual output							
To increase the library’s role as a viable partner in the research enterprise							
To reduce user dependence on your library’s print collection							
Other (Please specify in question 16b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don’t know, NA=Not applicable

16b. If you rated “Other’ for the question above, please specify in the box below.

17a. How would you assess your methods for recruiting digital content for the IR?

	VS*	SS	SU	VU	NO	DK	NA
Volunteer contributions							
Publicity about the IR in campus newspapers							
Presentations by staff responsible for the IR at departmental and faculty meetings							
Personal visits by staff responsible for the IR to faculty and administrators							
Staff responsible for the IR working one-on-one with early adopters							
Word-of-mouth from early adopters to their colleagues in the faculty and staff ranks							
Publicizing the IR during reference interactions in libraries and archives							
Systematic review of faculty, staff, center, and departmental web sites for potential contributors by staff responsible for the IR							

Institution-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints, etc.							
Other (Please specify in question 17b below)							

* Key to abbreviations: VS=Very successful, SS=Somewhat successful, SU=Somewhat unsuccessful, VU=Very unsuccessful, NO=No opinion, DK=Don't know, NA=Not applicable

17b. If you rated "Other" for the question above, please specify in the box below.

J. IR Implementation

18a. What IR software package have you implemented? (Choose one only.)

	Pilot Tested	Implemented
ARNO		
bePress		
CDSWare		
ContentDM		
DigiTool (Ex Libris)		
DiVA		
Documentum		
Dpubs		
DSpace		
Fedora		
GNU Eprints		
Greenstone		
HarvestRoad Hive		
Innovative Interfaces		
i-TOR		
Luna		
myCORE		
OPUS		
Sunsite		
Virginia Tech ETD software		
None		
Other (Please specify in question 18b below)		

18b. If you checked "Other" for the question above, please specify in the box below.

19. How would you characterize your IR's host? (Choose one only.)

A regional or state-based consortium
A partnership that joins your institution with one or more comparable institutions
Your institution only
A for-profit vendor

A not-for-profit vendor
Other (please specify)

20. What interoperability standards does your IR support? (Choose all that apply.)

IR supports OAI-MPH
IR is OpenURL compliant
IR materials use persistent identifiers
Our institution’s federated searching includes the IR
Other (please specify)

21a. Based on your experience with IR implementation, how would you rate your chosen system with regard to these capabilities?

	VA*	SA	SI	VI	NO	DK	NA
Technical support							
Technical documentation							
Adherence to open access standards							
Scalability = System growth and enhancement							
Customization							
Extensibility = Access to other campus systems and data							
Supported file formats							
User authentication							
Formulating metadata for digital documents							
Browsing, searching, and retrieving digital content							
End-user interface generally							
Controlled vocabulary searching							
Authority control							
Digital preservation							
Other (Please specify in question 21b below)							

* Key to abbreviations: VA=Very adequate, SA=Somewhat adequate, SI=Somewhat inadequate, VI=Very inadequate, NO=No opinion, DK=Don’t know, NA=Not applicable

21b. If you rated “Other” for the question above, please specify in the box below.

22a. If your efforts to implement an IR involved pilot testing IR software packages, what were the most important benefits of the pilot testing?

	VI*	SI	SU	VU	NO	DK	NA
Giving demonstrations to people involved in the IR implementation decision							
Giving demonstrations to an institution (s) interested in partnering with us to encourage them in IR implementation							
Gauging the interest of potential contributors to the IR							

Gauging the interest of potential IR-system users							
Identifying the strengths and shortcomings of available IR software							
Estimating costs for the technical implementation of an operational IR							
Developing the requisite technical expertise for IR implementation							
Identifying first adopters of an IR at your institution							
Control over your institution's intellectual output							
Preservation of your institution's intellectual output							
Other (Please specify in question 22b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

22b. If you rated "Other" for the question above, please specify in the box below.

23. If your efforts to implement an IR involved early adopters of IR technology, from what academic colleges, departments, schools, and service units have they come? (Choose all that apply.)

Your institution's library
Your institution's central computing unit
Your institution's archives
A particular academic college, department, or school
A particular service unit
Don't know
Not applicable
Other (please specify)

K. IR Content

24. Estimate the total number of digital documents that are published or in process in your IR.

25a. Estimate the number of digital documents that make up your IR's collections. (Write in the amount or write in DK for Don't Know or NA for Not Applicable.)

Preprints
Working papers
Books
Journals
Journal articles
Maps
Interview transcripts
Sound recordings of interview transcripts
Software
Software documentation
Video recordings of performances

Blogs
Interim and final reports to funding agencies
Raw data files that result from faculty research projects
Raw data files that result from doctoral dissertation research
Raw data files that result from master’s thesis research
Raw data files that result from senior thesis research
Written papers or transcripts of conference presentations
Conference presentations (e.g., summaries, abstracts, notes, outlines, remarks, etc.)
Committee meeting agenda and minutes
Committee meeting documents, e.g., budgets, reports, memoranda
Your institution’s course catalogs
Your institution’s newspapers
Your institution’s alumni publications
Faculty senate agendas and minutes
College, departmental, and school alumni publications
Regent, trustee, board meeting agendas and minutes
Course syllabi, class notes, handouts, outlines, assignments prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Other learning objects such as simulations, models, software demonstration files, images, video prepared by faculty, lecturers, teaching assistants, and other professional teaching personnel
Doctoral dissertations
Master’s theses
Senior theses
Graduate student eportfolios
Undergraduate student eportfolios
Class notes, outlines, assignments, papers, and projects prepared by graduate students
Class notes, outlines, assignments, papers, and projects prepared by undergraduate students
Other (Please specify type of digital document in question 25b below)

25b. If you entered an estimate for “Other” in the previous question, please specify in the box below.

26a. What file formats have you guaranteed contributors that you will preserve in perpetuity?

	Guaranteed	DK*	NO	NA
Plain Text UTF-8 (Unicode)				
Plain Text ANSI X3.4/ECMA-6/US-ASCII (7-bit)				
Plain Text ISO 8859-x (8-bit)				
Plain Text (all other encodings, including, but not limited to ISO 646 national variants)				
Rich text				
XML				
TeX				
LaTeX				
Postscript				

PDF				
PDF/A				
Microsoft Word				
Microsoft Excel				
Microsoft PowerPoint				
TIFF				
GIF				
JPEG				
PNG				
BMP				
Photo CD				
Photoshop				
AIFF				
Audio/Basic				
MPEG audio				
AAC_M4A				
Real Audio				
Windows Media Audio				
Wave				
AVI				
MPEG-1				
MPEG-2				
MPEG-4				
Windows Media Video				
Quicktime				
Other (Please specify file format in question 26b below)				

* Key to abbreviations: NO=No opinion, DK=Don't know, NA=Not applicable

26b. If you selected "Other" in the previous question, please specify in the box below.

L. IR Policies

27. Who is responsible for managing the IR's intellectual property rights? (Choose all that apply.)

Contributors' academic or service units
One chosen academic unit
One chosen service unit
IR staff
Library staff
Archives staff
Staff from the office of the chief information officer
A company to which our IR is outsourcing
Other (please specify)

28a. What is the status of these IR policies?

	NP*	D	I	DK	NA
Determining what is acceptable content					
Defining collections					
Determining who is authorized to make contributions to the IR					
Restricting access to IR content					
Acceptable file formats					
Identifying metadata formats and authorized metadata creators					
Charging for IR services					
Formulating a privacy policy for registered IR system users					
Licensing IR content					
Updating IR content					
Withdrawing IR content					
Providing access management services					
Preserving IR content					
Revising IR policies in the future					
Authorizing external contributors					
Intellectual property					
Other (Please specify in question 28b below)					

* Key to abbreviations: NP=No policy; D=Drafted; I=Implemented; DK=Don't know, NA=Not applicable

28b. If you rated "Other" for the question above, please specify in the box below.

M. IR Deployment

29a. To what extent do you think the following are likely to inhibit your ability to deploy a successful IR?

	VL*	SL	SU	VU	NO	DK	NA
Making members of your institution's learning community aware of the IR							
Contributors' lack of knowledge about how they can benefit from IRs							
Encouraging faculty to submit digital content to the IR							
Convincing faculty that the IR will not adversely affect the current publishing model							
Absence of campus-wide mandates regarding mandatory contribution of certain material types, e.g., doctoral dissertations, master's theses, faculty preprints, etc.							
Contributors' concerns about the difficulty using the IR system to contribute digital content to the IR							

Inability of contributors to formulate quality metadata							
Contributors' concerns about intellectual property rights for digital materials							
Inadequacy of the IR system's digital preservation capabilities							
Difficulties in long-term preservation of digital files							
Lack of on-campus technical expertise in IR systems							
Supporting all ongoing costs of an operational IR							
Competing for resources with other priorities, projects, and initiatives							
Other (Please specify in question 29b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

29b. If you rated "Other" for the question above, please specify in the box below.

N. Relationships

30. To what extent will an IR affect your institution's ability to build relationships between the IR and other on-campus repositories (e.g., archives, student services, library systems, digital asset management systems, electronic course management systems, digital libraries)?

A big positive effect
A moderate positive effect
No effect
A moderate negative effect
A big negative effect
A combination of positive and negative effects
Don't know
No opinion
Not applicable
Other (please specify)

O. Funding

31a. How likely is it that funding for your institution's implementation of an IR will come from these sources?

	VL*	SL	SU	VU	NO	DK	NA
Special initiative supported by your institution's central administration							
Special initiative supported by your institution's library							
Special initiative supported by your institution's central computer services							

Special initiative supported by your institution's archives							
Special initiative supported by academic colleges, departments, and schools							
Regular budget line item for your institution's central administration							
Regular budget line item for your institution's library							
Regular budget line item for your institution's central computer services							
Regular budget line item for your institution's archives							
Regular budget line item for academic colleges, departments, and schools							
Costs absorbed in routine operating costs of your institution's central administration							
Costs absorbed in routine operating costs of your institution's library							
Costs absorbed in routine operating costs of your institution's central computer services							
Costs absorbed in routine operating costs of your institution's archives							
Costs absorbed in routine operating costs of your institution's academic colleges, departments, and schools							
Grant awarded by an external source							
Grant awarded by an internal source							
Other (Please specify in question 31b below)							

* Key to abbreviations: VL=Very likely, SL=Somewhat likely, SU=Somewhat unlikely, VU=Very unlikely, NO=No opinion, DK=Don't know, NA=Not applicable

31b. If you rated "Other" for the question above, please specify in the box below.

32a. What percentage of your IR's annual budget is allocated to these categories? (Percentages must add up to 100%.)

	% Staff (including benefits)
	% Hardware acquisition
	% Hardware maintenance
	% Software acquisition
	% Software maintenance and updates
	% System backup
	% Vendor fees (for IRs hosted by an external vendor)
	% Other (Please specify in question 32b below)

32b. If you provided a percentage for "Other" for the question above, please specify in the box below.

P. Future Migration

33. How long do you think your institution will stick to this IR system before migrating to a new system? (Please enter number of years.)

34. How likely are you to modify your IR's software?

Very likely
Somewhat likely
Somewhat unlikely
Very unlikely
Don't know
No opinion
Not applicable

35a. What do you think will be the most important reasons for migrating to a new IR system?

	VI*	SI	SU	VU	NO	DK	NA
Greater capacity for handling preservation							
Friendlier user interface							
Advanced searching features							
Friendlier digital content submissions procedure							
Better tools for assisting contributors with metadata creation							
Around-the-clock technical support							
Greater versatility with the wide range of digital formats							
Greater opportunities for customization							
Greater versatility for linking to other campus systems and data							
Other (Please specify in question 35b below)							

* Key to abbreviations: VI=Very important, SI=Somewhat important, SU=Somewhat unimportant, VU=Very unimportant, NO=No opinion, DK=Don't know, NA=Not applicable

35b. If you rated "Other" for the question above, please specify in the box below.

36. What approaches have you used to date to assess your IR's success? (Choose all that apply.)

Tracking number of contributions
Tracking number of unique contributors
Tracking number of searches
Tracking number of users
Tracking number of unique users
Tracking number of queries
Conducting interviews with IR contributors
Conducting interviews with IR users
Surveying IR contributors

Surveying IR users
Other (please specify)

Q. Institutional Information

37. Please identify your position at your institution. (Choose one only.)

President or chancellor
Staff in the office of the president or chancellor
Vice president or provost
Staff in the office of the vice president or provost
Chief information officer
Staff in the office of the chief information officer
Archivist
Archives staff
Library director
Assistant director of library public services
Assistant director of library technical services
Assistant director of library information technology
Library staff
Other (please specify)

38. What is your connection to your institution’s IR?

39. Please identify your institution.

40. If your institution's IR is available to the general public, please give its web address(es):

R. Follow-up information

41. How can the MIRACLE Project assist you regarding IRs?

42. If you would be willing to volunteer for follow-up questions via phone or email, please add your name and email address and we will contact you in the near future:

Name
Email

Thank you! If you have questions, please message Soo Young Rieh (rieh@umich.edu) at the MIRACLE project. Thank you for your responses.

APPENDIX F

Literature Review

A short introduction featuring institutional repository (IR) definitions (Appendix F1) and early IR projects (Appendix F2) sets the stage for a literature review mirroring the topics on which MIRACLE Project questionnaires queried census respondents: (1) the extent of involvement with IRs at academic institutions (Appendix F3); (2) the people involved with IRs (Appendix F4); (3) the budget for an IR (Appendix F5); (4) important investigative activities leading up to IR implementation (Appendix F6); (7) IR-system software, content, and features (Appendix F7); (8) IR practices and policies (Appendix F8); (9) IR benefits and beneficiaries (Appendix F9); and (10) institutions where no IR planning has been done (Appendix F10).

F1 IR Definition and Attributes

Branin (2005) traces the origins of IRs to the knowledge management (KM) movement. He cites the impact of Peter Drucker's seminal *Harvard Business Review* paper "The Coming of the New Organization," which proposes that knowledge is the most important asset of the modern organization. Drucker does not define "knowledge" in this article but instead says that it "manifest[s] itself in various forms ... patents, trade secrets, operational routines, expertise inside the heads of employees" (Branin 2005, 237). Branin (2005, 238) sums up by saying the "KM movement in the 1990s established the nomenclature of IRs, lit a fire under organizations to pay attention to knowledge assets, and forced organizations to define knowledge assets broadly."

The IR may also be a technology-based solution to the problem of escalating serials costs and reform of the scholarly communication process. Abstract models, guidance, and technology standards that have sparked the development of IRs and ensured their interoperability with other online systems are the Open Archival Information System (OAIS) Reference Model, the Open Access Initiative's Protocol for Metadata Harvesting (OAI-PMH), and the Metadata Encoding and Transmission Standard (METS) (Branin 2005).

Not unlike IRs are open-access e-print and preprint repositories that scientific and scholarly disciplines have promulgated for the dissemination of research within disciplines and subdisciplines. Instead of the discipline-based focus that underlines e-print and preprint repositories, IRs feature the intellectual output of educational and research units in universities, research institutes, and state and federal governments, and thus are more likely to be encyclopedic in

subject coverage, representing the full range of academic fields of study, especially IRs at universities, and including a wide range of document genres, e.g., working papers, simulations, specimens, and course syllabi.

Clifford Lynch (2003) defines IRs and enumerates their attributes:

An institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution ... It represents a collaboration amongst librarians, information technologists, archives and records managers, faculty, university administrators, policy makers. An institutional repository will be supported by a set of information technologies, but a key part of the services ... is the management of technological changes, and the migration of digital content from one set of technologies to the next as part of the organizational commitment to providing repository services.

F2 The First IR Systems, IR Projects, and IR-related Projects

Until there was a critical mass of institutions with operational IRs to survey, researchers occupied themselves by describing the first IR systems, IR projects, and IR-related projects. Examining open-access activity in the United Kingdom (UK), Pinfield (2003) focused on the FAIR (Focus on Access to Institutional Resources) Programme of the Joint Information Systems Committee (JISC). Shearer (2003a, 2003b) published periodic updates on IR efforts of libraries in the Canadian Association of Research Libraries (CARL). A report on the first 45 IRs and IR-related projects included an examination of publishers' responses to an online questionnaire about their attitudes toward IRs (Mark Ware Consulting 2004; Ware 2004a).

A key step forward for IRs in the United States was funding from The Andrew W. Mellon Foundation for the joint development of the open-source DSpace IR-software system by the Massachusetts Institute of Technology (MIT) Libraries and Hewlett-Packard Company (Smith et al. 2003). Since 2003, institutions with the requisite technical expertise have been able to download DSpace and customize it as needed for experimentation, pilot testing, or public implementation.

Another IR-system software package that has benefited from extensive foundation funding is Fedora (Flexible Extensible Digital Object and Repository Architecture). Fedora's origins can be traced to Defense Advanced Research Projects Agency (DARPA) and National Science Foundation (NSF) funding at Cornell University in 1997. Developed jointly by Cornell University Information Science and

the University of Virginia Library with grants from The Andrew W. Mellon Foundation, Fedora is a general-purpose digital repository service. Fedora applications are IRs, library collections management, multimedia authoring systems, archival repositories, and digital repositories (Fedora Project 2005–2006). Institutions that download Fedora's open-source software will need technical expertise to customize it.

F3 IR Surveys and the Extent of Survey Respondents' Involvement with IRs

In 2005, researchers began surveying educational institutions regarding their efforts with IRs. Lynch and Lippincott (2005) distributed their survey to the 124 member institutions of the Coalition for Networked Information (CNI) and to 81 liberal arts colleges bearing CNI consortial membership. They queried respondents about the extent of IR implementation at their institutions, database size, content, system software packages, policies, and administrative responsibility for the IR. Response rates were high—78.2% for CNI members and 43.8% for CNI consortial members.

About 40% of CNI members have “some type” of IR operating, and 88% of those who do not have an operational IR have undertaken IR planning. Only 6% of CNI-consortial members have an operational IR. Lynch and Lippincott (2005) note that “deployment of institutional repositories beyond the doctoral research institutions in the United States is extremely limited ... Most of the engagement ... is at colleges and universities where students and faculty have strong commitments to locally created materials for teaching and learning or that document student research.” They suggest two factors that might diversify the types of institutions involved with IRs: (1) large-scale student contributions in the form of electronic portfolios (e-portfolios) to IRs; and (2) a groundswell of faculty demand for IR services to facilitate the dissemination of their research and teaching objects.

Conference organizers from CNI, the JISC in the UK, and SURF Foundation in the Netherlands conducted an international survey to determine the current state of IR deployment in the academic sector (van Westrienen and Lynch 2005). Survey results reveal the extent of IR implementation by country, database size and content, system software packages, contributor activity, policies, and factors that stimulate and inhibit IRs. The researchers estimate “a spread from around 5% in a country like Finland, where repositories are just getting started, to essentially 100% deployment in countries like Germany, Norway and the Netherlands, where it is clear that repositories have already achieved some status as common infrastructure across the relevant national higher education sector and, hence, can form the basis for other initiatives that presuppose the near-universal availability of institutional repositories.”

Shearer's (2004) informal telephone survey of member libraries in the CARL reveals the extent of their involvement with IRs, system

software packages, content, recruiting contributors, and working with early adopters. Of the 20 universities she contacted, 25% have operational IRs and 35% are in the process of implementation. The remaining 40% are in the planning stages.

A team of University of Houston researchers surveyed the 123 member libraries of the Association of Research Libraries (ARL) regarding their IR efforts and published findings in an ARL SPEC Kit (Bailey et al. 2006). Completing the survey were 87 (71%) ARL libraries. Survey content was comprehensive; it queried ARL members about the extent of their involvement with IRs, people involved in the IR effort, budgets, policies, recruiting content, system software packages, benefits, and evaluating IRs. Of the 87 respondents, 37 (43%) have implemented IRs, 31 (35%) are planning for IRs, and 19 (22%) have no plans for IR involvement. The survey researchers note that perceptions of staff planning for an IR do not always correspond with experiences of staff who have implemented an IR, particularly concerning time and resources required to implement an IR and the level of difficulty in recruiting content.

F4 The People Involved with IRs

At the heart of Lynch's (2003) definition of an IR is a "collaboration of librarians, information technologists, archives and records managers, faculty, university administrators, and policy makers." Pelizzari (2005) advocates the library as "the standard bearer for the IR." Quint (2002, 8) urges librarians to assert themselves as the leaders in IR implementation. She writes, "This is the best chance librarians will ever have to break the chains that have bound them and their budgets ... Who will step up and help to create a better process of scholarly communications? ... If academic librarians do not step up to pay that price and right now, they could find themselves blocked out of that future, and perhaps, any future at all. Now or never." A survey of CNI members and consortial members reports that 80% of institutions with operational an IR have put sole responsibility for the IR into the hands of the library.

The ARL SPEC Kit survey results reveal that the library plays a critical role in initiating, planning, and implementing IRs (Bailey et al. 2006). "All respondents, implementers and planners alike, indicate that the library has been a driving force in the creation of or planning for an IR" (Bailey 2006, 14). Following a successful IR implementation, the library department that manages digital library initiatives usually takes charge of IR operations and collaborates with other library departments such as technical services, archives, and/or cataloging. The ARL SPEC Kit study reports IR staffing unusually high numbers (Bailey et al. 2006, 15):

If the mean FTE [full-time equivalent] values for each of the four units are added together, the average number of staff working on an implementer's 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 planner's 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.

Gibbons (2004, 17) advocates a partnership of librarians, instructional technologies training staff, and computer services staff for the IR effort. Ware (2004a) asserts that IRs start with a partnership between the library, the institution's instructional technology unit, and a vice president's office (e.g., bursar or provost for academic studies).

IR staff at institutions where IRs have been implemented have been eager to tell their story (e.g., Phillips, Carr, and Teal 2005; Jones, Andrew, and MacColl 2006, 159–189; Chan, Kwok, and Yip 2005; Rogers 2003). The IR effort usually involves central administrative, academic, and service units. For example, Hewlett-Packard Labs and MIT software engineers, MIT librarians, faculty, and administrators, and early adopters were involved in the development and deployment of DSpace (Baudoin and Branschofsky 2003). Case studies generally present a range of experiences such as IR-system software development, system implementation and deployment, policy creation, collection development strategies, and preservation initiatives.

Librarians have been enthusiastic about carving out a role for themselves vis-à-vis the IR. Allard, Mack, and Feltner-Reichert (2005) identify librarians in these roles: learning how the IR works; leading the IR implementation effort; developing policy (especially with regard to defining collections); leading anything that pertains to metadata, reviewing submissions to the IR; and training contributors. Additional roles include evaluating IR systems, being an advocate for the IR, recruiting content, and serving as advisory contributors on intellectual property issues (Chan, Kwok, and Yip 2005).

Archivists are more wary of their relationships with the IR. Crow (2002a) observes this phenomenon.

Depending on the university, an institutional repository may complement or compete with the role served by the university archives. University archives often serve two purposes: (1) to manage administrative records to satisfy legally mandated retention requirements, and (2) to preserve materials pertaining to the institution's history and to the activities and achievements of its officers, faculty, staff, students, and alumni. Compared to institutional repositories, which aim to preserve the entire intellectual output of the institution, university archivists exercise broad discretion in determining which papers and other digital objects to collect and store. Still, the potential overlap of roles of the two repository types merits consideration at institutions that support both.

Crow's custodial and passive view flies in the face of recent efforts of archives to become more involved in the research and teaching missions of colleges and universities. He also downplays the difficulties in documenting all the intellectual output of an institution and the benefits of a selection process in the IR collection develop-

ment plan. Arguing for selection of faculty output, Bicknese (2003–2004) posits that archivists bring to the table substantial expertise in collecting and appraising faculty papers as well as other university records and recommends that IRs use this expertise.

F5 The Budget for an IR

The single largest line item in the budget for an IR is staffing (Gibbons 2004, 54). Technical staff, at the very least, a systems administrator and a programmer, are needed to profile, program, maintain and customize IR-software systems. Jones, Andrew, and MacColl (2006) advocate an IR staff composed of systems developer, a liaison officer, a project manager, and a metadata editor. Librarians should be responsible for recruiting content from faculty, students, and other members of an academic institution's learning community. Costs for software, equipment, and backup equipment are minor compared with staffing costs.

Ballpark figures for annual costs of the IR effort are \$285,000 at MIT (Barton and Walker 2002), \$265,000 at The Ohio State University (OSU) (Branin 2005, 247), and \$200,000 at the University of Rochester (Gibbons 2004, 56). The ARL Spec Kit reports these costs: (1) for institutions planning an IR, a median of \$75,000 and a range from \$12,000 to \$160,000; (2) for institutions implementing an IR, a median of \$45,000 and startup costs ranging from \$8,000 to \$1,800,000; and (3) for institutions that have implemented an IR, a median of \$41,750 and operations costs ranging from \$8,600 to \$500,000 (Bailey et al. 2006). "The primary method of funding both start-up and ongoing costs is reallocation from existing budgets" (Bailey et al. 2006, 16).

At MIT, Barton and Walker (2002) break down costs into three major line items: (1) staff salaries and benefits, (2) operating expenses, and (3) system equipment escrow. Although MIT is unique because of its development of the DSpace IR, its funding model may inform other institutions contemplating local development. The MIT model anticipates financial support and in-kind assistance from four sources: (1) the MIT Libraries' operating budget, (2) collaborative development of DSpace that results from related projects that deploy DSpace, (3) enhancement of DSpace by DSpace Federation members at whose institutions DSpace is deployed, and (4) charging for premium services.

"One of the likely largest costs over the long term will be ... preservation ... also by far the least known and indeed least knowable ... [s]o a commitment to an IR amounts to an implicit commitment to an unknown amount of work at some point in the future" (Mark Ware Consulting 2004, 24). Strategies for recovering costs include depositing funds into escrow accounts in anticipation of the future costs of preservation, hosting digital content submitted by partner institutions, and charging for ancillary services such as file conversion, digitizing physical artifacts, metadata creation, and exceeding file-storage quotas (Gibbons 2004, 56; Barton and Walker 2002). Branin (2005, 247) asks rhetorically, "What will it cost an insti-

tution or society not to provide stewardship of its important digital assets?”

F6 Important IR Investigative Activities

Institutions beginning an IR project have two comprehensive sources to consult. From the *LEADIRS Workbook* (Barton and Waters 2004–2005), they will learn about planning an IR, choosing IR-system software, developing policies, anticipating costs, and recruiting content. The workbook features worksheets that institutions can use to guide them through IR planning and implementation, and it links to online examples of planning and implementation documents, e.g., service definitions, promotional brochures, needs assessment surveys, operational IRs, and policy statements. The second resource, *The Institutional Repository*, provides “a comprehensive outline of the main issues to consider when setting-up and developing an institutional repository—from making the case within the institution and choosing suitable software to formulating workflows, policy, and advocacy plans” (Jones, Andrew, and MacColl 2006, xvi).

Investigative activities that IR committees conducted prior to implementing OSU’s Knowledge Bank IR include a data-warehouse planning project to identify administrative data needed for decision support, an environmental scan of operational and in-development IRs, discussions with representative faculty groups, an inventory of on-campus digital initiatives, a review of relevant technology standards, and a pilot-project compilation of faculty publications (Rogers 2003; Baudoin and Branschovsky 2003; Walters 2006).

F7 IR-system Software, Content, and Features

Gibbons’ (2004) article in *Library Technology Reports* is a comprehensive examination of IR systems and system features. Institutions approaching pilot-test and implementation phases of the IR effort would benefit from a periodic update of her article, especially now that several for-profit vendors have entered or are about to enter the marketplace with new IR systems. Barton and Waters’ (2004–2005) *LEADIRS Workbook* also lists available IR software systems and uniform resource locators (URLs) that link to systems in operation. Smith et al. (2003) and Tansley et al. (2003) restrict themselves to the design and functionality of the open-source DSpace IR.

Because those who responded to the CNI survey count IR content in different ways, Lynch and Lippincott (2005) do not make an effort to estimate the size of operational IRs. The types of materials CNI members submit to their IRs run the gamut, representing the research, teaching, service, and publicity activities of colleges and universities, and ranging from text-based items to multimedia artifacts—images, audio, video, software, blogs, and e-portfolios. The majority (58%) of CNI respondents have implemented DSpace; other popular systems are bepress, ContentDM, Virginia Tech’s ETD, DigiTool, and locally developed systems. Most CARL institutions are

using DSpace (Shearer 2004). CARL IRs feature a wide variety of artifact types, but the rate of deposit is slow and sporadic (Shearer 2004). DSpace is also the most commonly used IR-system software by ARL-member libraries that are planning IRs or operating IRs (Bailey et al. 2006). Some ARL libraries that have implemented an IR are using DSpace in conjunction with other IR-system software, such as ContentDM or ETD-db. The few ARL implementers who are not DSpace institutions have chosen ContentDM, ProQuest Digital Commons, and Archimède.

Content in the IRs of the 13 nations participating in the international survey sponsored by CNI, JISC, and SURF is divided between journal articles and theses; the sole exception is Australia, where 88% of content is primary data (van Westrienen and Lynch 2005). Institutions in 11 of the 13 countries are using DSpace, and institutions in 7 of the 13 countries are using GNU ePrints software. Institutions in Germany favor OPUS while those in Australia favor Virginia Tech's ETD. Discussing the size of Open Archives Initiative (OAI) repositories, ePrints archives, and IRs, Mark Ware Consulting (2004, 33) concludes that "the majority [of IRs] are clearly in a very early stage of development but even more of the longer-established sites have a relatively small number of documents compared to the research outputs of their institutions."

F8 IR Practices and Policies

Devising IR policies is an important, necessary, and complex activity during IR implementation. A survey of CNI members and consortial members reports that 60% of institutions with an operational IR put sole responsibility for IR policy making in the hands of the library. Laundry lists of policy issues are given by Mark Ware Consulting (2004), Barton and Waters (2004–2005), and Shearer (2005). In preparation for the Library and Information Technology Association (LITA) Regional Institute "Establishing an Institutional Repository," Gibbons (2005) compiled a list of links to the IR policies of nearly 20 universities worldwide. Another source of online IR policy data is a section of the ePrints Web site called ROARMAP (Registry of Open Access Repository Material Archiving Policies) that invites institutions to record their commitment to providing open access to peer-reviewed research and to share their departmental or institutional policies regarding open-access provision (ePrints, n.d).

F8.1 Digital Content for the IR

Gibbons (2004, 21–26), Lynch (2003), and Crow (2002a) describe potential digital content for the IR that results from faculty who are active researchers and teachers (e.g., e-prints, working papers, technical reports, conference presentations, data sets, and learning objects) and students who do research and want to document their academic accomplishments (e.g., theses, dissertations, data sets, and e-portfolios). IRs can be exacting about acceptable types of materials. For example, MIT's DSpace allows only publisher-accepted materials, and

the University of California's eScholarship accepts only materials that authors have not submitted to publishers (Shearer 2002).

F8.2 Making Contributions to the IR

Contributing content involves metadata capture, file management, and license handling (Jones, Andrew, and MacColl 2006). Policies can specify acceptable metadata formats, required and optional metadata tags, and people who are authorized to submit, review, and update metadata. File management pertains to checking file formats, data integrity of files, and the completeness of documents.

Copyright infringement is one of the important challenges that IR staff face. When submitting manuscripts to publishers, most academics sign copyright-transfer agreement forms that assign copyright to publishers. Surveying 542 academic authors, Gadd, Oppenheim, and Proberts (2003) report that 49% reluctantly assign copyright to publishers and 41% do so freely. Surveying 1,200 authors who publish in medicine and the life sciences, Hoorn and van der Graaf (2006) report that 71% prefer to keep copyright, 2% prefer to transfer copyright to publishers, 23% are neutral about the choice between the two, and 4% do not know. A little less than half of their survey's respondents want to keep copyright, allow unlimited reuse of published works for scholarly and educational purposes, and put limitations on reuse for commercial purposes.

So many authors assign copyright to publishers that, for the time being, IR staff may have to be wary of IR contributors infringing on copyright and advise them about copyright issues. If authors assign the copyright for their works to publishers, IR staff should advise authors to scrutinize their copyright agreements with publishers to make sure they retain the right to publish the work or the last pre-publication version electronically (Barton and Waters 2004–2005).

F8.3 Preservation

Long-term preservation of scholarly content is an essential role of IRs (Lynch 2003). Not a one-time event, long-term preservation consists of specific functions such as ingesting digital objects in which metadata is created, storing such objects and associated metadata, monitoring technology obsolescence, and evaluating the usage of digital objects (Fyffe et al. 2004).

At the moment, IRs have not become the equivalent of trusted digital repositories "whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now, and in the future" (RLG 2002). Trusted digital repositories are responsible for long-term maintenance of digital objects with organizational, technical and financial sustainability (Jantz and Giarlo 2005). The RLG-OCLC conception of trusted digital repositories requires a certification process, but the degree to which libraries would pursue this type of certification is not known. Certification would impose a very different type of accountability—an *external* accountability—on libraries, something that they have not previously experienced (RLG 2002).

Many IRs employ shared standards such as the Open Archival Information System (OAIS) reference model and the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The OAIS model provides a comprehensive framework of all the functions required for digital preservation including ingest, storage, retrieval, and long-term preservation of digital objects. OAI-PMH is a mechanism for harvesting extensible markup language (XML)-based metadata from repositories, and it therefore makes possible interoperable search and retrieval among repositories (Branin 2005). Implementation of digital preservation in IRs, however, is still in its infancy.

Preservation concerns of libraries and institutions of higher education are also shared by IR contributors and users. Swan and Brown (2004a, 2004b) note that 42% of open-access (OA) journal authors are worried about preservation of OA journal contents. Rowlands et al. (2004), however, state that their survey's respondents disagreed with the idea that OA journals are ephemeral. Although the concern about preservation "should not be a good reason for eschewing open access (indeed, all online only) journals" (Swan and Brown 2004b, 67), preservation of open-access contents, as well as the larger issue of institutional repository contents, is an important issue that needs to be investigated in greater depth (Kim 2006).

Although no consensus on best practices exists, digital preservation has these key functional goals: (1) data can be maintained over time without being lost, damaged, or altered; (2) data can be found and served to users; (3) data can be interpreted and understood by users; and (4) goals 1, 2, and 3 can be fulfilled in a long-term plan (Wheatley 2004).

From a policy standpoint of preservation, identifying file formats for which IRs provide long-term preservation is necessary. When making decisions about preserving file formats, Jones, Andrew, and MacColl (2006, 80) suggest IR staff answer these questions: (1) Is the file format an open standard/format? (2) Is the file format widely used? (3) Is the file format and associated technology likely to be preserved? (4) Are the contents of the file human readable? and (5) Is the file format itself human readable? These authors assert that preserving formats to which IR staff answer positively to questions 4 and 5 is best.

The ARL-sponsored survey reports that 74% of ARL libraries' operational IRs accept any digital file type. "Relatively few (26%) are committed to functional preservation of every file type" (Bailey et al. 2006).

F8.4 Recruiting IR Content

Recruiting content for IRs has been difficult. Except in the Netherlands and Germany, the number of academics contributing to or knowledgeable of IRs is very low (van Westrienen and Lynch 2005). A year after the IR at the University of Toronto Scarborough became operational, contributors had submitted hardly 1,000 items to it (Chan 2004, 287). Ware's (2004b) survey of 45 early IRs reports an average number of documents of 1,256; he qualifies the largest IR

at the University of Virginia (21,000 items) because two-thirds of its contents come from a digital photograph collection. “Difficult” sums up the response of about two-thirds (63%) of ARL libraries regarding the recruitment of digital content for the IR (Bailey et al. 2006).

Examples of strategies that IR staff have enlisted to boost IR submission rates are hiring work-study students to do proxy archiving for faculty, giving presentations at faculty meetings, demonstrating operational IRs, working with early adopters and relying on them to tell their colleagues and students about the IR, long-term archiving of locally edited e-journals, and encouraging local Web site developers to relocate digital content from their sites to the IR (Crow 2002b, 23; Gibbons 2004, 57; Chan 2004; Shearer 2004, 2005; Mackie 2004; Barton and Waters 2004–2005; Jenkins, Breakstone, and Hixson 2005; Graham, Skaggs, and Stevens 2005; Chan, Kwok, and Yip 2005; Bell, Foster, and Gibbons 2005; Bailey et al. 2006). Because the evidence in favor of higher citation rates for open-access material is mounting (Odlyzko 2000; Lawrence 2001; Swan and Brown 2004a; Antelman 2004; Eysenbach 2006a, 2006b), scholars may be inclined to bypass traditional publication channels and publish instead in open-access repositories in the not-so-distant future.

On the basis of the results of an ethnographic study, Foster and Gibbons (2005) take a radically different approach to recruiting IR content from faculty at the University of Rochester. They reconfigure DSpace so that faculty contributions take place in a personal Web page or curriculum vitae (CV) that features vital information about contributors such as name, title, and contact information, and archives in the IR their research and teaching output (e.g., course syllabi, conference presentations, working papers, preprints).

F9 IR Benefits and Beneficiaries

Some IR advocates envision IRs as the critical component in reforming the existing system of scholarly communication. Researchers would have 24–7 access to scholarship, full-text search and retrieval capacity, and the ability to link backward and forward in time using footnotes and citations (Harnad 2001b). Academic institutions would reap benefits such as the following (Crow 2002a, 2002b; Chan 2004):

- a new scholarly publishing paradigm that wrests control from publishers and puts it back in the hands of the academy
- increased visibility, prestige, and public value
- maximal access to the results of publicly funded research
- increased numbers and diversity of scholarly materials that are collected and preserved

In the ARL Spec Kit Survey (Bailey et al. 2006), the top three reasons ARL libraries give for implementing IRs are institution-centered—to increase global visibility of, to preserve, and to provide free access to the institution’s scholarship.

Many barriers prevent authors from fully embracing self-archiving. Examples are authors’ lack of knowledge about self-ar-

chiving, questions about whether self-archiving will erode the quality control that results from peer review and will make authors' works more susceptible to plagiarism, the time required to submit documents to IRs, technical difficulties that authors encounter while making contributions, the possibility of infringing on signed copyright agreements, authors' concerns about the equivalence of their institution's IR and a journal publisher, distrust of their institution's commitment to long-term maintenance, and a reluctance to tamper with the current working system of scholarly communication (Harnad 2001a; Chan 2004; Pinfield 2004, 2005; Swan and Brown 2004a, 2004b; Swan 2005b; Ober 2005; Jenkins, Breakstone, and Hixson 2005; van Westrienen and Lynch 2005).

Pelizzari (2005) voices his concern: "The biggest obstacle may be inertia ... amongst academics ... [the problems of] intellectual property rights, quality control, workload (their own), undermining the 'tried and tested' publishing status quo on which academic reputations and promotions lie." About 70% to 80% of authors admit that they would comply with mandates from employers or funders regarding submission to an open-access archive (Swan 2005b; Swan and Brown 2004b). Harnad et al. (2004) are vocal advocates for mandates or incentives from employers and funders. Carr and Harnad (2005) report the results of studying two months of submissions into a mature repository, estimate the amount of time spent entering metadata would be as little as 40 minutes per year for a highly active researcher, and conclude that self-archiving is not a time-consuming task for authors or their designees.

In the ARL SPEC Kit survey (Bailey et al. 2006), most IR implementers perceive content recruitment to be difficult; half of the planners are neutral and the other half is divided between "easy" and "difficult." This finding reveals a change in perception regarding content recruitment between IR planning and IR implementation. IR implementers employ especially diverse and aggressive content-recruitment strategies, digitizing print materials and depositing them in the IR for authors and holding symposia to raise awareness of IRs among prospective authors.

F10 Institutions Where No IR Planning Has Been Done

To date, most of the information we have about IRs targets IR implementation especially at large research institutions. Little is known about institutions that are in the IR preplanning or planning stages or about the experiences of institutions that have a teaching focus. CNI, CNI-JISC-SURF, CARL, and ARL surveys report on the numbers and percentages of such institutions (see Appendix F3), but they report little else about them.

REFERENCES

Allard, Suzie, Thura R. Mack, and Melanie Feltner-Reichert. 2005. The Librarian's Role in Institutional Repositories: A Content Analysis of the Literature. *Reference Services Review* 33(3):325–336.

American Library Association. 2006. ALA Library Fact Sheet 3—Lists of Libraries. Available from <http://www.ala.org/Template.cfm?Section=libraryfactsheet&Template=/ContentManagement/ContentDisplay.cfm&ContentID=23385>.

Antelman, Kristin. 2004. Do Open-Access Articles Have a Greater Research Impact? *College & Research Libraries* (September):372–382. Available from http://eprints.rclis.org/archive/00002309/01/do_open_access_CRL.pdf.

Bailey, Charles W. Jr., Karen Coombs, Jill Emery, Anne Mitchell, Chris Morris, Spencer Simons, and Robert Wright. 2006. *Institutional Repositories*. Washington, D.C.: Association of Research Libraries. SPEC Kit 292

Barton, Mary R., and Julie Harford Walker. 2002 (July). MIT Libraries' DSpace Business Plan Project: Final Report to the Andrew W. Mellon Foundation. Available from <http://www.dspace.org/implementation/mellon.pdf#search=%22barton%20mellon%22>

Barton, Mary R., and Margaret M. Waters. 2004–2005. *Creating an Institutional Repository: LEADIRS Workbook*. Cambridge, Mass.: Massachusetts Institute of Technology Libraries. Available from <http://dspace.org/implementation/leadirs.pdf>.

Baudoin, Patsy, and Margret Branschofsky. 2003. Implementing an Institutional Repository: The DSpace Experience at MIT. *Science & Technology Libraries* 24(1/2):31–45.

Bell, Suzanne, Nancy Fried Foster, and Susan Gibbons. 2005. Reference Librarians and the Success of Institutional Repositories. *Reference Services Review* 33(3):283–290.

Bicknese, Douglas. 2003-2004. Institutional Repositories and the Institution's Repository: What Is the Role of the University Archives with an Institution's On-line Digital Repository? *Archival Issues* 28(2):81–93.

Borgman, Christine L. 2006. *Disciplines, Documents, and Data: Convergence and Divergence in the Scholarly Information Infrastructure*. Institute for Scientific Information's Samuel Lazerow Memorial Lecture, University of Tennessee, October 18. Available from <http://www.sis.utk.edu/newsArchive/lazerow2006>.

Branin, Joseph. 2005. Institutional Repositories. In Miriam A. Drake, ed. *Encyclopedia of Library and Information Science*. Boca Raton, Fla.: Taylor & Francis Group, LLC. Draft available from https://kb.osu.edu/dspace/bitstream/1811/441/1/inst_repos.pdf.

The Carnegie Foundation for the Advancement of Teaching. 2006a. The Carnegie Classification of Institutions of Higher Education. Available from <http://www.carnegiefoundation.org/classifications/index.asp>.

———. 2006b. Carnegie Classifications: Lookup & Listings. Available from <http://www.carnegiefoundation.org/classifications/index.asp?key=807>.

Carr, Leslie, and Stevan Harnad. 2005 (March 15). Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving. Available from <http://eprints.ecs.soton.ac.uk/10688/01/KeystrokeCosting-publicdraft1.pdf>.

Chan, Diana L. H., Catherine S. Y. Kwok, and Steve K. F. Yip. 2005. Changing Roles of Reference Librarians: The Case of the HKUST Institutional Repository. *Reference Services Review* 33(3):268–282.

Chan, Leslie. 2004. Supporting and Enhancing Scholarship in the Digital Age: The Role of Open-Access Institutional Repositories. *Canadian Journal of Communication* 29:277–300. Available from http://eprints.rclis.org/archive/00002590/01/Chan_CJC_IR.pdf.

Creative Commons. 2006. Enabling the Legal Sharing and Reuse of Cultural, Educational, and Scientific Works. Available from <http://creativecommons.org/>.

Crow, Raym. 2002a. The Case for Institutional Repositories: A SPARC Position Paper. Available from <http://www.arl.org/sparc/IR/ir.html>.

———. 2002b. SPARC Institutional Repository Checklist & Resource Guide. Available from http://www.arl.org/sparc/IR/IR_Guide_v1.pdf.

Drabenstott, Karen M. 1994. *Analytical Review of the Library of the Future*. Washington, D.C.: Council on Library Resources.

ePrints. n.d. ROARMAP (Registry of Open Access Repository Material Archiving Policies). Available from <http://www.eprints.org/openaccess/policysignup/>.

Eysenbach, Gunther. 2006a. Citation Advantage of Open Access Articles. *PLoS Biology* 4(5) e157:0692–0698. Available from http://biology.plosjournals.org/archive/1545-7885/4/5/pdf/10.1371_journal.pbio.0040157-L.pdf.

———. 2006b. The Open Access Advantage. *Journal of Medical Internet Research* 8(2): e8. Available from <http://www.jmir.org/2006/2/e8>.

Fedora Project. 2005–2006. Fedora. Available from <http://www.fedora.info/index.shtml>.

Foster, Nancy Fried, and Susan Gibbons. 2005. Understanding Faculty to Improve Content Recruitment for Institutional Repositories. *D-Lib Magazine* 11(1). Available from <http://www.dlib.org/dlib/january05/foster/01foster.html>.

Fyffe, R., D. Ludwig, M. Roach, B. Schulte, and B. F. Warner. 2004. Preservation Planning for Digital Information: Final Report of the HVC² Digital Preservation Task Force. University of Kansas. Available from <https://kuscholarworks.ku.edu/dspace/bitstream/1808/166/1/Preservation+Planning+for+Digital+Information.pdf>.

Gadd, Elizabeth, Charles Oppenheim, and Steve Proberts. 2003. The Intellectual Property Rights Issues Facing Self-Archiving. *D-Lib Magazine* 9(9). Available from <http://www.dlib.org/dlib/september03/gadd/09gadd.html>.

Gibbons, Susan. 2004. Establishing an Institutional Repository. *Library Technology Reports* 40(4):11–14.

Gibbons, Susan. 2005. Institutional Repository Policies on the World Wide Web Prepared for LITA Regional Institute “Establishing an Institutional Repository,” October 27, 2005. Available from <http://docushare.lib.rochester.edu/docushare/dsweb/Get/Document-19828/Policies.pdf>.

Graham, John-Bauer, Bethany Latham Skaggs, and Kimberly Weatherford Stevens. 2005. Digitizing a Gap: A State-Wide Institutional Repository Project. *Reference Services Review* 33(3):337–345.

Harnad, Stevan. 2001a. The Budapest Open Access Initiative Self-Archiving FAQ. Available from <http://eprints.ecs.soton.ac.uk/10635/01/index.html>.

———. 2001b. For Whom the Gate Tolls: How and Why to Free the Refereed Research Literature Online through Author/Institution Self-Archiving, Now. Available from <http://www.ukoln.ac.uk/events/open-archives/presentations/harnad.pdf>.

Harnad, Stevan. 2006. Harnad Eprint Archive and Psycology and BBS Journal Archives. Available from <http://www.ecs.soton.ac.uk/~harnad/>.

Harnad, Stevan, Tim Brody, François Vallières, Les Carr, Steve Hitchcock, Yves Gingras, Charles Oppenheim, Heinrich Stamerjohanns, and Eberhard R. Hilf. 2004. The Access/Impact Problem and the Green and Gold Roads to Open Access. *Serials Review* 30(4):310–314.

Hoorn, Esther, and Maurits van der Graaf. 2006. Copyright Issues in Open Access Research Journals. *D-Lib Magazine* 12(2). Available from <http://www.dlib.org/dlib/february06/vandergraaf/02vandergraaf.html>.

ICPSR (Inter-University Consortium for Political and Social Research). 2005. *Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle*. Ann Arbor, Mich.: Institute for Social Research. Available from <http://www.icpsr.umich.edu/access/dataprep.pdf>.

Jantz, Ronald, and Giarlo, Michael J. 2005. Digital Preservation: Architecture and Technology for Trusted Digital Repositories. *D-Lib Magazine* 11(6). Available from <http://www.dlib.org/dlib/june05/jantz/06jantz.html>.

Jenkins, Barbara, Elizabeth Breakstone, and Carol Hixson. 2005. Content in, Content Out: The Dual Roles of the Reference Librarian in Institutional Repositories. *Reference Services Review* 33(3):312–324.

Jones, Richard, Theo Andrew, and John MacColl. 2006. *The Institutional Repository*. Oxford, U.K.: Chandos Publishing.

Kim, Jihyun. 2006. Motivating and Impeding Factors Affecting Faculty Contribution to Institutional Repositories. Paper presented at the Joint Conference on Digital Libraries 2006 Workshop: Digital Curation & Trusted Repositories: Seeking Success, Chapel Hill, N.C., June 15, 2006. Available from <http://sils.unc.edu/events/2006jcdl/digitalcuration/Kim-JCDLWorkshop2006.pdf>.

Lawrence, Steve. 2001. Free Online Availability Substantially Increases a Paper's Impact. *Nature* 411 (May 31):521. Available from <http://www.nature.com/nature/journal/v411/n6837/pdf/411521a0.pdf>.

Lynch, Clifford A. 2003. Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL Bimonthly Report* 226:1–7. Available from <http://www.arl.org/newsltr/226/ir.html>.

Lynch, Clifford A., and Joan K. Lippincott. 2005. Institutional Repository Deployment in the United States as of Early 2005. *D-Lib Magazine* 11(9). Available from <http://www.dlib.org/dlib/september05/lynch/09lynch.html>.

Lynch, Clifford A. 1992. Reaction, Response, and Realization: From the Crisis in Scholarly Communication to the Age of Networked Information. *Serials Review* 18(1/2): 107–112.

Mackie, Morag. 2004. Filling Institutional Repositories: Practical Strategies from the Daedulus Project. *Ariadne* 39. Available from <http://www.ariadne.ac.uk/issue39/mackie/>.

Mann, Thomas. 1993. *Library Resource Models*. New York: Oxford University Press.

Mark Ware Consulting Ltd. 2004. Publisher and Library/Learning Solutions (PALS): Pathfinder Research on Web-Based Repositories: Final Report. Available from [http://www.palsgroup.org.uk/palsweb/palsweb.nsf/0/8c43ce800a9c67cd80256e370051e88a/\\$FILE/PALS%20report%20on%20Institutional%20Repositories.pdf](http://www.palsgroup.org.uk/palsweb/palsweb.nsf/0/8c43ce800a9c67cd80256e370051e88a/$FILE/PALS%20report%20on%20Institutional%20Repositories.pdf).

Markey, Karen. 2007. The Online Library Catalog: Paradise Lost or Paradise Regained? *D-Lib Magazine* 13 (1/2). Available from <http://www.dlib.org/dlib/january07/markey/01markey.html>.

Ober, John L. 2005 (March 31). Postprint Repository Services: Context and Feasibility at the University of California. Available from http://osc.universityofcalifornia.edu/responses/materials/UC_postprint-study_final.pdf.

Odlyzko, Andrew. 2000 (March 19). The Rapid Evolution of Scholarly Communication. Available from <http://www.si.umich.edu/PEAK-2000/odlyzko.pdf>.

Pelizzari, Eugenio. 2005. Opinions, Research, and Analysis: Harvesting for Disseminating: Open Archives and the Role of Academic Libraries. *The Acquisitions Librarian* 17(33/34):35–51.

Phillips, Holly, Richard Carr, and Janis Teal. 2005. Leading Roles for Reference Librarians in Institutional Repositories: One Library's Experience. *Reference Services Review* 33(3):301–311.

Pinfield, Stephen. 2001. How Do Physicists Use an E-Print Archive? *D-Lib Magazine* 7(12). Available from <http://www.dlib.org/dlib/december01/pinfield/12pinfield.html>.

———. 2003. Open Archives and UK Institutions: An Overview. *D-Lib Magazine* 9(3). Available from <http://www.dlib.org/dlib/march03/pinfield/03pinfield.html>.

———. 2004. Self-Archiving Publications. In G. E. Gorman and Rowland Fytton, eds. *International Yearbook of Library and Information Management 2004-2005: Scholarly Publishing in an Electronic Era*. London: Facet. Available from <http://eprints.nottingham.ac.uk/archive/00000142/01/IYLIM04.PDF>.

———. 2005. A Mandate to Self Archive? The Role of Open Access Institutional Repositories. *Serials* 18(1) (March):30–34.

Queensland University of Technology. 2006. Policy F/1.3 E-print Repository for Research Output at QUT. November 3. Available from http://www.mopp.qut.edu.au/F/F_01_03.html.

Quint, Barbara. 2002. Now or Never! *Searcher* 10(1):6, 8.

RLG. 2002. *Trusted Digital Repositories: Attributes and Responsibilities*. Mountain View, Calif.: RLG, Inc. Available from <http://www.rlg.org/en/pdfs/repositories.pdf>.

Rogers, S. A. 2003. Developing an Institutional Knowledge Bank at Ohio State University: From Concept to Action Plan. *portal: Libraries and the Academy* 3(1):125–136.

RoMEO SHERPA. 2006. Opening Access to Research. Available from <http://www.sherpa.ac.uk/index.html>.

Rosenberg, Victor. 1966. *The Application of Psychometric Techniques to Determine the Attitudes of Individuals toward Information Seeking*. Bethlehem, Pa.: Lehigh University. AD 637 713.

Rowlands, Ian, David Nicholas, and Paul Huntington. 2004. Scholarly Communication in the Digital Environment: What Do Authors Want? *Learned Publishing* 17(4):261–273.

Shearer, Kathleen. 2002 (September). A Guide to Setting-up an Institutional Repository. Available from http://www.carl-abrc.ca/projects/institutional_repositories/setup_guide-e.html.

———. 2003a. The CARL Institutional Repositories Pilot Project. *D-Lib Magazine* 9(3). Available from <http://www.dlib.org/dlib/march03/03inbrief.html>.

———. 2003b. Institutional Repositories: Towards the Identification of Critical Success Factors. *Canadian Journal of Information and Library Science* 27(3):89–108.

———. 2004 (September). Survey Results—Summer 2004: CARL Institutional Repositories Project. Available from http://www.carl-abrc.ca/projects/institutional_repositories/pdf/survey_results_2004-e.pdf.

———. 2005 (April 28). Institutional Repositories: The Evolution of Scholarly Communication. Available from http://www.carl-abrc.ca/projects/institutional_repositories/ppt/CACUL2-Apr05.ppt.

Smith, MacKenzie, Mary Barton, Mick Bass, Margret Branschofsky, Greg McClellan, Dave Stuve, Robert Tansley, and Julie Harford Walker. 2003. DSpace: An Open Source Dynamic Digital Repository. *D-Lib Magazine* 9(1). Available from <http://www.dlib.org/dlib/january03/smith/01smith.html>.

Suber, Peter. 1996–2006. Peter Suber. Available from <http://www.earlham.edu/~peters/hometoc.htm>.

Swan, Alma. 2005a (April). Open Access: Briefing Paper. JISC. Available from http://www.jisc.ac.uk/uploaded_documents/JISC-BP-OpenAccess-v1-final.pdf.

———. 2005b (May). Open Access Self-Archiving: An Introduction. Key Perspectives Limited. Available from <http://eprints.ecs.soton.ac.uk/11006/01/jiscsum.pdf>.

Swan, Alma, and Sheridan Brown. 2004a. Authors and Open Access Publishing. *Learned Publishing* 17(3):219–224.

Swan, Alma P., and Sheridan N. Brown. 2004b (February). JISC/OSI Journal Authors Survey Report. Key Perspectives Ltd. Available from http://www.jisc.ac.uk/uploaded_documents/JISCOAreport1.pdf.

Tansley, Robert, Mick Bass, David Stuve, Margret Branschofsky, Daniel Chudnov, Greg McClellan, and MacKenzie Smith. 2003. The DSpace Institutional Digital Repository System: Current Functionality. *Proceedings of the 2003 Joint Conference on Digital Libraries (JCDL '03)*. Available from <https://dspace.mit.edu/handle/1721.1/26705>.

U.S. Department of Education. 2006. Title III Part A Programs: Strengthening Institutions. Available from <http://www.ed.gov/programs/iduestitle3a/index.html>.

van Westrienen, Gerard, and Clifford A. Lynch. 2005. Academic Institutional Repositories: Deployment Status in 13 Nations as of Mid-2005. *D-Lib Magazine* 11(9). Available from <http://www.dlib.org/dlib/september05/westrienen/09westrienen.html>.

Walters, Tyler O. 2006. Strategies and Frameworks for Institutional Repositories and the New Support Infrastructure for Scholarly Communications. *D-Lib Magazine* 12(10). Available from <http://www.dlib.org/dlib/october06/walters/10walters.html>.

Ware, Mark. 2004a. Institutional Repositories and Scholarly Publishing. *Learned Publishing* 17(2):115–124.

Ware, Mark. 2004b. Universities' Own Electronic Repositories yet to Impact on Open Access. *Nature*. Available from <http://www.nature.com/nature/focus/accessdebate/4.html>.

Wheatley, Paul. 2004. Institutional Repositories in the Context of Digital Preservation. Available from <http://www.dpconline.org/docs/DPCTWf4word.pdf>.